6-8-2017

Towards Automated Analysis of Urban Infrastructure after Natural Disasters using Remote Sensing

Colin Axel
cma4204@rit.edu

Towards Automated Analysis of Urban Infrastructure after Natural Disasters using Remote Sensing

by

Colin Axel

B.S. Rochester Institute of Technology, 2013

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy
in the Chester F. Carlson Center for Imaging Science
College of Science
Rochester Institute of Technology

June 8, 2017

Signature of the Author _____

Accepted by _____
Coordinator, Ph.D. Degree Program Date

CHESTER F. CARLSON CENTER FOR IMAGING SCIENCE
COLLEGE OF SCIENCE
ROCHESTER INSTITUTE OF TECHNOLOGY
ROCHESTER, NEW YORK

CERTIFICATE OF APPROVAL

Ph.D. DEGREE DISSERTATION

The Ph.D. Degree Dissertation of Colin Axel
has been examined and approved by the
dissertation committee as satisfactory for the
dissertation required for the
Ph.D. degree in Imaging Science

Dr. Jan van Aardt, Dissertation Advisor

Dr. David Ross, External Chair

Dr. David Messinger

Dr. Carl Salvaggio

Date

Disclaimer

The views, opinions, findings and conclusions reflected in this publication are solely those of the author and do not represent the official policy or position of the USDOT/OST-R, or any State or other entity. USDOT/OST-R does not endorse any third party products or services that may be included in this publication or associated materials.

Towards Automated Analysis of Urban Infrastructure after Natural Disasters using Remote Sensing

by

Colin Axel

Submitted to the
Chester F. Carlson Center for Imaging Science
in partial fulfillment of the requirements
for the Doctor of Philosophy Degree
at the Rochester Institute of Technology

Abstract

Natural disasters, such as earthquakes and hurricanes, are an unpreventable component of the complex and changing environment we live in. Continued research and advancement in disaster mitigation through prediction of and preparation for impacts have undoubtedly saved many lives and prevented significant amounts of damage, but it is inevitable that some events will cause destruction and loss of life due to their sheer magnitude and proximity to built-up areas. Consequently, development of effective and efficient disaster response methodologies is a research topic of great interest.

A successful emergency response is dependent on a comprehensive understanding of the scenario at hand. It is crucial to assess the state of the infrastructure and transportation network, so that resources can be allocated efficiently. Obstructions to the roadways are one of the biggest inhibitors to effective emergency response. To this end, airborne and satellite remote sensing platforms have been used extensively to collect overhead imagery and other types of data in the event of a natural disaster. The ability of these platforms to rapidly probe large areas is ideal in a situation where a timely response could result in saving lives. Typically, imagery is delivered to emergency management officials who then visually inspect it to determine where roads are obstructed and buildings have collapsed. Manual interpretation of imagery is a slow process and is limited by the quality of the imagery and what the human eye can perceive.

In order to overcome the time and resource limitations of manual interpretation, this dissertation investigated the feasibility of performing fully automated post-disaster analysis of roadways and buildings using airborne remote sensing data. First, a novel algorithm for detecting roadway debris piles from airborne light detection and ranging (lidar) point clouds and estimating their volumes is presented. Next, a method for detecting roadway flooding in aerial imagery and estimating the depth of the water using digital elevation models (DEMs) is introduced. Finally, a technique for assessing building damage from airborne lidar point clouds is presented. All three methods are demonstrated using remotely sensed data that were collected in the wake of recent natural disasters.

The research presented in this dissertation builds a case for the use of automatic, algorithmic analysis of road networks and buildings after a disaster. By reducing the latency between the disaster and the delivery of damage maps needed to make executive decisions about resource allocation and performing search and rescue missions, significant loss reductions could be achieved.

Contents

List of Figures

List of Tables

List of Acronyms

2-D	two-dimensional. 7, 9, 28, 30, 34, 67, 68, 85
3-D	three-dimensional. xv, 7, 11, 12, 15, 18, 24, 28, 50, 67, 68, 71, 76, 115
ALS	airborne laser scanning. 11
ARP	access restoration plan. 102
AVIRIS	Airborne Visible / Infrared Imaging Spectrometer. 10
DEM	digital elevation model. vii, 2, 4, 9, 23, 28, 49, 50, 55, 57, 60, 61, 64–66, 75, 99, 100
DNDWI	Difference of Normalized Difference Water Indices. 46
DSM	digital surface model. 17, 18, 69
EMR	electromagnetic radiation. 9, 10
EMS	electromagnetic spectrum. 9
EMS-98	European Macroseismic Scale. 85, 87
FIS	fuzzy inference system. 16
FN	false negative. 30, 32, 59, 85–88
FP	false positive. 30, 32, 59, 85–88
FPA	focal plane array. 8, 9, 15
GCS	geographic coordinate system. 9, 11
GIS	geographic information system. 20, 22, 30, 43, 85
GLCM	gray-level co-occurrence matrix. 47
GPS	global positioning system. 8, 9, 11
GSD	ground sample distance. 46, 50, 114
HAZUS	Hazards United States. 14
HSV	hue, saturation, and value. 52
IMU	inertial measurement unit. 8, 9, 11
lidar	light detection and ranging. vii, xi, 2–4, 7, 8, 11, 12, 15–24, 26, 32–36, 39, 42, 43, 47, 49, 50, 68, 69, 71, 72, 74–76, 90, 92, 98–100, 102, 108, 109, 111, 114, 120
LOS	line of sight. 8
MWIR	midwave infrared. 47
nDSM	normalized digital surface model. 17, 23, 75, 80, 87

NDWI Normalized Difference Water Index. 48, 65
NIR near-infrared. 9, 46–48, 50, 52, 59–61, 64, 65, 100
NYC New York City. 108–110

PMF Progressive Morphological Filter. 23, 74, 75

QSF quadratic spectral filter. 55, 65

RGB red, green, and blue. 50, 52, 59–61, 63, 64, 100, 114, 115
ROC receiver operating characteristic. 54

SAM spectral angle mapper. 48, 53–55, 65
SAR Synthetic Aperture Radar. 45, 49, 67, 68
SFM structure from motion. 114, 115, 120
SVM Support Vector Machine. 69
SWIR shortwave infrared. 9, 10, 48

TN true negative. 87
TP true positive. 30, 32, 59, 85–88

UAS unmanned aerial system. xii, 100, 114, 115, 117, 120, 121
UAV unmanned aerial vehicle. 8, 15, 46, 47, 65
USA United States of America. 50, 66, 108
UTM Universal Transverse Mercator. 33

VHR very high resolution. 17

WASP Wildfire Airborne Sensor Program. 20, 30, 72, 79, 81, 82, 84, 85, 88, 96

List of Institutions

3DEP	three-dimensional (3-D) Elevation Program. 50, 64
CRED	Centre for Research on the Epidemiology of Disasters. 5
EC JRC	European Commission Joint Research Centre. 85, 94
FEMA	Federal Emergency Management Agency. 110
MOTF	Modeling Task Force. 110
NGS	National Geodetic Survey. 50, 59, 66
RIT	Rochester Institute of Technology. 19, 72
UNITAR	United Nations Institute for Training and Research. 19
UNOSAT	UNITAR Operational Satellite Applications Program. 19, 85, 94
WB	World Bank. 85, 94

List of Symbols

Σ	covariance matrix. 53
$\mathbf{\Sigma_b}$	covariance matrix of background pixels. 54, 55
$\mathbf{\Sigma_s}$	covariance matrix of reference pixels. 54, 55
α	The alpha radius for computing alpha shapes [m]. 28
γ	standard deviation multiplier for statistical outlier removal. 23, 74
$\mathbf{s}$	reference spectral vector. 53, 54
$\mathbf{x}$	data spectral vector. 54, 55
μ	mean. 53
$\boldsymbol{\mu_b}$	the mean spectral vector for a distribution of background pixels. 54, 55
μ_k	mean of all the mean distances between a point and its k neighbors for the entire point cloud. 22, 23, 28, 74
$\boldsymbol{\mu_s}$	the mean spectral vector for a distribution of reference pixels. 54, 55
σ_f	standard deviation of the local variation of surface normals. 77
σ_k	standard deviation of all the mean distances between a point and its k neighbors for the entire point cloud. 22, 23, 74
σ_v	standard deviation of curvature. 77
C	Indicator function for buildings that are candidates for damage assessment. 80
β	The angle between the normal vectors of two lidar points [^{o}]. 26, 77
c	The speed of light in vacuum: $\approx 3 \cdot 10^8$ [m/s]. 11
D	data term for graph cuts. 76, 77
d	Euclidean distance [m]. 24, 76
Dam	damage classification for buildings. 83
$\bar{d}$	mean Euclidean distance [m]. 22, 23, 74
δ	the indicator function for graph cuts. 77
dh_0	initial elevation difference threshold for the Progressive Morphological Filter [m]. 74
E	energy function for graph cuts. 76
e_s	water surface elevation. 57
f	variation of surface normals. 76, 77

$\mathcal{G}$	the weighted graph of points used for graph cuts. 76
h	height above ground [m]. 25, 26, 80, 82, 83
h_w	water height [m]. 57
HR	height ratio of an object. 27
I	Input flood image. 52
I_1	index of lidar points classified as non-debris due to low height. 26
I_2	index of lidar points classified as non-debris due to smoothness. 26
k	number of neighbors to a point to include in neighborhood operations. 22, 74
l	label of a point in a weighted graph. 76, 77
λ	eigenvalue of point neighborhood covariance matrix. 24, 76
λ^n	eigenvalue of normal vector neighborhood covariance matrix. 25, 26, 76
$\vec{n}$	normal vector. 24, 25, 76, 77
N_p	local neighborhood of lidar points. 24, 25, 75, 76
P	point cloud. 23, 24, 26, 74, 76
p	lidar point. 22–24, 26, 74–77
P_o	the lidar points that make up an object produced from region growing. 27
P^*	point cloud resulting from statistical outlier removal. 23, 74
q	all of the lidar points in a point cloud. 24, 76, 77
R	range [m]. 11
r	radius [m]. 24, 75–78
R_s	ratio of segmented points to unsegmented points. 83
r_{QSF}	detection statistic for QSF. 55
r_{SAM}	detection statistic for SAM. 54
S	smoothness term for graph cuts. 76, 77
s_f	scalar coefficient to weight the normal variation term in the data term for graph cuts. 77

s_s scalar coefficient that weights the smoothness term for graph cuts. 77

s_v scalar coefficient to weight the curvature term in the data term for graph cuts. 77

t The time of flight to the target and back [ns]. 11

T_β Threshold on the angle between the normal vectors of two points [^{o}]. 26, 78, 83

T_d Threshold on the ratio of highly angled points to total points of a building. 80

T_h Threshold on the height above ground [m]. 25–27, 80

T_{h2} Threshold on the height above ground, used for damage classification [m]. 83

T_{HR} Threshold on the height ratio of an object. 27

T_l Threshold on the ratio of low points to total points in a building, used for candidate identification. 80

T_{l2} Threshold on the ratio of low points to total points in a building, used for damage classification. 83

T_λ Threshold on the maximum variation of normals on the Gaussian sphere. 25, 26

T_p Minimum threshold on the number of points in a building region. 78

T_s Threshold on the ratio of segmented points to total points in a building. 83

T_θ Threshold on the angle between the normal vector and the horizontal [^{o}]. 26, 80

T_v Threshold on the curvature of a point during region growing. 78, 83

T_{QSF} Threshold on the QSF detection statistic image. 55

T_{SAM} Threshold on the SAM detection statistic image. 54

θ The angle between a normal vector and the horizontal [^{o}]. 24–26, 80, 81

$\mathcal{V}$ 3-D points in the graph used for graph cuts. 76

v curvature. 76–78

Chapter 1

Introduction

1.1 Context

Natural disasters, such as earthquakes and hurricanes, are an unpreventable component of the complex and changing environment we live in. Continued research and advancement in disaster mitigation through prediction and preparation have undoubtedly saved many lives and prevented immense amounts of damage, but it is inevitable that some events will cause destruction and loss of life due to their sheer magnitude and proximity to built-up areas. Consequently, development of effective and efficient disaster response methodologies is a research topic of great interest.

A successful emergency response is dependent on a comprehensive understanding of the scenario at hand. It is crucial to assess the state of the infrastructure and transportation network so that resources can be allocated efficiently. Obstructions to the roadway are one of the biggest inhibitors to effective emergency response. To this end, airborne and satellite remote sensing platforms have been used extensively to collect overhead imagery and other types of data in the event of a natural disaster. The ability of these platforms to rapidly probe large areas is ideal in a situation where a timely response could result in saving lives. Typically, imagery is delivered to emergency management officials who then visually inspect it to determine

where roads are blocked and buildings have collapsed. Manual interpretation of imagery is a slow process (a disaster over a single city could require hundreds or thousands of images for complete coverage) and is limited by the quality of the imagery and what the human eye can perceive.

In the interest of streamlining the disaster response, a vast amount of research has focused on automatically or semi-automatically processing remotely sensed data to generate information products such as maps of road obstructions (*e.g.*, building debris, floodwater, *etc.*) and damaged buildings. In most studies, imagery or lidar data are used to determine the location of a region or object of interest. Although the object locations are extremely helpful to response coordinators, they could construct a more efficient plan of action if the disaster features were quantified in terms of size and severity (*e.g.*, the volume of a debris pile). This dissertation advances automated analysis of post-disaster infrastructure using remotely sensed data in terms of three distinct objectives.

1.2 Objectives

This dissertation investigates the use of remotely sensed data to perform fully automatic analysis of the road network and infrastructure of a region recently impacted by a natural disaster. The dissertation is structured in terms of three primary objectives:

1. Assess the ability to use airborne lidar point clouds to automatically detect individual piles of debris in the roadway and estimate their volume.

   1.1. Determine the completeness, correctness, and quality of debris detection using debris piles manually digitized in reference imagery.

   1.2. Quantify the error in estimating debris volume as a function of point density using reference volume data obtained through an experiment.

2. Evaluate the ability to use multispectral, aerial imagery and DEM rasters to automatically detect roadway flooding and estimate the inundation depth.

2.1. Calculate the classification accuracy of the proposed algorithm using floodwaters manually digitized in reference imagery.

3. Assess the ability to use airborne lidar point clouds to automatically detect buildings and classify them as damaged or undamaged.

3.1. Calculate the damage classification accuracy using ground truth damage assessments.

1.3 Layout

This dissertation is composed of six chapters. A brief description of each chapter (not including the introduction) is as follows:

1.3.1 Chapter 2: Background

The background chapter starts off with a broad view of natural disasters and the disaster management cycle and then focuses on the use of remote sensing in natural disasters. A discussion of airborne remote sensing as a whole is presented and then digital imaging and lidar are described in detail.

1.3.2 Chapter 3: Roadway Debris Detection and Volume Estimation using Airborne Lidar

This chapter proposes an automated workflow to detect debris piles in the road network using airborne lidar point clouds and estimate their volume. Local surface features and region growing are utilized to extract roadway debris piles and then the piles are reconstructed as surfaces using alpha shapes in order to estimate their volume. Results are presented using lidar data collected after the 2010 earthquake in Haiti. An additional refinement focuses on how lidar density impacts our assessment.

1.3.3 Chapter 4: Detection and Depth Estimation of Roadway Flooding using Airborne Imagery and Digital Elevation Models

Chapter 4 introduces an algorithm for automatically detecting flooded roadways and estimating the depth of the water in the road using airborne imagery and DEMs. Signature-matched detection is used to detect flood pixels in the image after a seed flood pixel is selected through entropy analysis of the input image. The depth of the water is calculated using the elevation of the land at the land-water border. An evaluation of the algorithm performance is presented using imagery from the severe flooding of the Mississippi River in January 2016.

1.3.4 Chapter 5: Building Damage Assessment using Airborne Lidar

In Chapter 5 an end-to-end, automated workflow for classifying buildings as damaged or undamaged using airborne lidar point clouds is proposed. Local surface properties and region growing are used to cluster the points into individual buildings and then features such as surface normal angle, planarity, and height above ground are used to make a classification of damaged or undamaged. Results from testing the algorithm on airborne lidar point clouds collected after the 2010 earthquake in Haiti are presented.

1.3.5 Chapter 6: Conclusions

Chapter 6 wraps up the dissertation with a summary of the work presented throughout the main research chapters (3-5) and draws some conclusions based on the project as a whole.

Chapter 2

Background

2.1 Natural Disasters

Natural disasters are an inevitable threat to humanity. Geophysical (*e.g.*, earthquakes and volcanic activity), hydrological (*e.g.*, floods and landslides), and meteorological (*e.g.*, storms) disasters are three examples of the types of natural disasters that occur annually across the globe (Guha-Sapir *et al.*, 2015). Natural disasters have significant physical and economic impacts on society. According to the Centre for Research on the Epidemiology of Disasters (CRED), natural disasters are responsible for 99,820 deaths, $162.5B in economic losses, and affect 199.2 million people on average each year (Guha-Sapir *et al.*, 2015). Moreover, the frequency of these types of events have increased threefold from a global annual average of approximately 100 in the 1980's, to greater than 300 in present day (Kousky, 2014).

To help prepare for and cope with natural disasters, emergency planners use the disaster management cycle, which contains four phases of both planning and action: reduction, readiness, response, and recovery (Karen *et al.*, 2009). Reduction focuses on decreasing the likelihood of a disaster occurring through the use of structural installations (*e.g.*, stopbanks and levees) and land use planning. Identifying areas that are too dangerous to build in, ensuring that communities have ample routes for access and response, and

fortifying infrastructure are all examples of how planning can reduce the risk of a hazard becoming a disaster. Readiness, like reduction, is a planning stage in which measures are taken to establish efficient protocols in the case of a disaster. Common readiness practices include educating the public on disasters and evacuation procedures, training exercises for emergency responders, and installing early alert systems. In most cases, risk reduction and readiness are insufficient to prevent natural disasters. The response phase is crucial to minimize the impacts of the disaster (and even prevent secondary disasters) through damage and needs assessment, as well as prioritization and allocation of resources for making roads accessible to search and rescue missions, and evacuation. Finally, Recovery consists of both short term action (*i.e.*, restoration of communication and utilities) and long term action (*i.e.*, reconstruction of buildings and communities). All four phases of the disaster management cycle are important and should be integrated by emergency planners for the best possible results. However, emergency response is typically the most highly funded and planned stage, in part because of its complexity, but also because it is inevitable that many natural hazards will evolve into disasters, despite risk reduction.

The emergency response to a natural disaster is comprised of several logistical tasks and needs to be initiated immediately to prevent further losses and secondary hazards (*e.g.*, a chemical leak caused by an earthquake). Typical response tasks include: damage assessment, needs assessment, response prioritization, and deployment of resources (Mehrotra *et al.*, 2003). Damage assessment involves identifying affected buildings, locating obstructions to the transportation network, and estimating the severity of these impacts. During the needs assessment phase, circumstances requiring a response are determined, *e.g.*, survivors trapped under a collapsed building, earthquake debris blocking the evacuation route. The needs that require response are compared with available resources in the prioritization phase and then resources are allocated accordingly. Lastly, the response decisions are shared with emergency responders, who then utilize the resources to address the needs. The response follows a logical order, and when executed in a timely fashion can be instrumental in preventing unnecessary loss. In order to assess damage and needs, data need to be collected over the entire extent of the disaster region. However, most disasters are large enough in spatial scale that complete coverage through field surveys is impossible. The most common solution to providing

a comprehensive overview of the scenario is to collect remote sensing imagery, either by commissioning an aerial survey or utilizing an existing satellite within range to capture the scene.

Remote sensing imagery is well-suited for emergency response because it provides a rapid means of assessing the location and severity of the disaster impact and identifying areas in which the transport network or other critical facilities have been obstructed or destroyed. Many different types of imagery, and other data, can be acquired using remote sensing, and the appropriate modality should be determined based on specific spatial, spectral, temporal, and radiometric requirements (Karen *et al.*, 2009). For example, if one needs to monitor the heating and cooling cycles of a volcano, then thermal imagery could be used. As mentioned above, identifying and assessing the damage done to infrastructure is one of the primary tasks during the emergency response. High spatial resolution, optical imagery is often collected to locate collapsed buildings, as well as identify damage and obstructions in the transportation network. The level of detail present in high resolution imagery makes it simple to interpret for human analysts and a good data source for automated image processing routines. In some cases, the two-dimensional (2-D) information available in images is insufficient for determining how many resources must be sent out to resolve a problem caused by a disaster. Imagine that an image taken after an earthquake clearly shows a building collapsed into the main road that lies between a hospital and the epicenter of the earthquake. An emergency planner would then have to either: i) make a guess as to how much debris needs to be removed and then send resources to the scene, or ii) send a crew out to the scene to perform a field survey of the debris and then deploy the correct resources. If airborne lidar data were also collected, the volume of the debris could be estimated directly from the data, eliminating the need to send out a field crew or risk sending not enough machinery and slowing down the operation. This is only one of the many examples of why collecting a 3-D representation of the scene can be beneficial to disaster response. Although there are many forms of remote sensing data that are useful at various stages of the disaster management cycle, this dissertation focuses on the use of high resolution aerial imagery and airborne lidar data for emergency response.

2.2 Airborne Remote Sensing

In the planning stages of a remote sensing mission, the choice of platform and sensor is one of the most crucial decisions to be made. Spatial resolution, extent of coverage, spectral range, and temporal requirements are amongst the driving factors that will ultimately lead to a final choice. Both airborne (*i.e.*, from an aircraft) and spaceborne (*i.e.*, from a satellite) remote sensing have advantages in the context of emergency response and damage/debris assessment. Airborne systems can operate at low altitudes, thus providing fine spatial detail that can ensure the accuracy of output information products. Satellite-based systems typically provide broader coverage in a single frame, and don't need to be mobilized to begin collecting imagery. However, the flexibility of airborne systems in terms of spatial resolution, as well as increased control of imaging parameters and temporal scheduling, make them a better candidate for immediate collection of data following a disaster, when compared to commercial satellite systems that have limited spatial and temporal coverage (Stow *et al.*, 2015). Additionally, lidar data from commercial surveying companies are only available via aircraft at this time. Given the reasons presented in this section, most post-disaster remote sensing missions are tasked to airborne platforms.

In this dissertation, only remote sensing data collected from aircraft are utilized. This section provides a brief background on airborne imaging and lidar systems and their applications.

2.2.1 Digital Imaging

The collection of aerial imagery is a remote sensing practice that has been around from as early as 1858, when Gaspard Felix Tournachon experimented with photography from a hot air balloon using wet plates (Schott, 2007). Airborne imaging has come a long way since then; most imaging takes place on airplanes or unmanned aerial vehicles (UAVs) and images are captured through the use of electro-optical imaging systems. A typical aircraft-mounted system consists of several subsystems, including, but not limited to: focusing optics with pointing components, a focal plane array (FPA), readout electronics, a global positioning system (GPS), and an inertial measurement unit (IMU). The pointing system directs the line of sight (LOS) of the imager to the

particular region on the ground that is being imaged. The optics focus incoming photons onto the detector elements that make up the FPA. Readout electronics are responsible for filtering, amplifying and digitizing the analog signals from the FPA, ultimately producing a digital image (Eismann, 2012). The GPS and IMU record the exact position and orientation of the sensor as an image is collected, respectively. Leveraging the GPS and IMU information, images can be projected into a geographic coordinate system (GCS) and resampled to a grid in which each pixel has an equal scale in a process known as georectification. If a DEM is also available, variations in elevation can be corrected for by projecting the image into an orthorectified space, thus removing the effects of image perspective and relief (Schott, 2007).

One important characteristic of aerial imagery is the number of spectral bands it contains. Each spectral band is a sampling of the incident electromagnetic radiation (EMR) across a specific range of wavelengths. The four main types of imagery include: panchromatic, color, multispectral, and hyperspectral. Panchromatic imagery consists of a single band and is the result of a detector integrating photons over one broad bandwidth, typically spanning the visible portion of the electromagnetic spectrum (EMS) and sometimes into the near-infrared (NIR). For example, the IKONOS satellite has a panchromatic band over the 0.45-0.90μm spectral range (Schott, 2007). Panchromatic images therefore are displayed in black and white, because they only contain one spectral band.

Color images have three bands (red, green, and blue) and are sometimes referred to as true color, because they closely match what humans sees due to the three types of photoreceptors in the eye that have peak sensitivities in the red, green, and blue. Depending on the type of sensor, color images can be formed using multiple linear arrays with appropriate spectral filters, a single 2-D array with filter overlays, or a spectral beam splitter directing light to three identical 2-D arrays (Messinger, 2014).

Multispectral imagery refers to images with about four to ten relatively broad, non-contiguous bands whose central wavelengths are generally aligned with the application of the relevant imaging system (Messinger, 2014). Some examples of popular multispectral bands are: red, green, blue, red edge (0.7μm), and various NIR and shortwave infrared (SWIR) wavelengths. Multispectral images are typically generated by stacking filtered linear arrays in the along- or cross-track direction, depending on what type of imaging system is

being used. One example of a multispectral imaging system is the WorldView-2 satellite, which contains one broad, panchromatic band from 0.45-0.8μm and eight, narrower spectral bands ranging from 0.40 to 1.05μm (Messinger, 2014).

Hyperspectral images consist of many bands, typically more than 100, sampling narrow and contiguous bandwidths across the visible and SWIR wavelengths. The high number of spectral bands is achieved by splitting broadband EMR through a dispersive element, such as a prism or diffraction grating. Due to the increased number of spectral bands, hyperspectral imaging systems often have coarser spatial resolution than panchromatic or multispectral systems to keep the cost and size reasonable (Shaw and Burke, 2003). The Airborne Visible / Infrared Imaging Spectrometer (AVIRIS) system is a good example of an airborne hyperspectral imager. It captures 224 contiguous spectral bands between 0.4-2.5μm at a bandwidth of 10nm and a spatial resolution of approximately 20m (Schott, 2007).

The four types of imagery described in this section all have characteristics that make them useful for different applications. Panchromatic imagery typically has a high spatial resolution compared to other bands in the same imaging system, making it useful for spatial detection applications and enhancing lower spatial resolution spectral bands (captured simultaneously) through a process known as pan-sharpening (Padwick *et al.*, 2010; Abraham and Sasikumar, 2013). Multispectral imagery, and to a lesser extent, color imagery, is useful for visualization and classification of different land cover classes through the use of spatial and spectral features. Hyperspectral imagery has the advantage of producing fine spectral signatures for each pixel, making it ideal for target detection and material identification tasks, but at a coarser spatial resolution than afforded by multispectral imagery. The best type of digital imagery for disaster response is somewhat dependent on the task, but multispectral, or even color imagery, is most commonly used (Stow *et al.*, 2015). The spatial resolution that can be achieved by airborne multispectral imagers (typically finer than 1m) allows obstructed roadways, collapsed buildings, and other post-disaster features to be readily observed by human analysts or algorithms. Having three or more spectral bands allows for spectral or combined spatial-spectral classification of pixels into different classes, including post-hazard classes, such as floodwater or building debris.

2.2.2 Lidar

Airborne lidar, sometimes referred to as airborne laser scanning (ALS), is an active remote sensing modality that uses a ranging sensor mounted on an airborne platform, a scanning mechanism, and orientation and position sensors to provide a georeferenced, 3-D sampling of the targets below. The ranging sensor is typically comprised of an emitting laser (pulsed or continuous wave) and a receiver (Wehr and Lohr, 1999). A single laser pulse will reflect off of one or more targets, producing peaks in the backscattered energy profile collected by the receiver. Depending on the type of system, significant peaks will be digitized as discrete returns (discrete-return lidar) or the entire waveform will be recorded (waveform lidar) (Popescu, 2011). In discrete return lidar, each digitization produces a point in the main data product, the point cloud, which is a collection of 3-D points. For this dissertation, only lidar data produced from discrete-return, pulsed laser systems are used, and therefore the remaining background information is limited to those systems. The range, R, from the sensor to the reflection surface is calculated using the speed of light, c, and the time-of-flight of the laser pulse, t (equation 2.1).

$$R = \frac{c \cdot t}{2} \tag{2.1}$$

An opto-mechanical scanner, such as an oscillating mirror, is used to rapidly move the laser in equal angle steps, sampling the targets as points in the cross-track direction. The forward motion of the platform generates the point cloud samples in the along-track direction. For every pulse of the laser, a GPS measures the precise position of the sensor and an IMU provides the orientation of the sensor. The position and orientation information, coupled with the range measurement, allows each point in the point cloud to be assigned a 3-D coordinate in a GCS. Additionally, most modern systems are capable of recording an intensity value for each point, which is the maximum of the returned pulse or signal integration over the returned pulse width (Wehr and Lohr, 1999).

Airborne lidar is considered to be one of the most important geospatial data acquisition technologies as its inherent data structure provides information about both the elevation and the 3-D structure of the

underlying surfaces sampled by the point cloud (Popescu, 2011). Many fields of research, spanning both environmental and industrial domains, have utilized the 3-D power of airborne lidar data to improve or create novel methodologies. One common application is city mapping, in which buildings are extracted and reconstructed with accurate dimensions from lidar point clouds. In the case of a disaster, such as an earthquake, the same processing can be applied, and taken a step further, to detect and categorize damage to buildings and other infrastructure. A thorough discussion of the use of airborne lidar for natural disasters will be presented in chapters 3 and 5.

Chapter 3

Roadway Debris Detection and Volume Estimation using Airborne Lidar

3.1 Introduction

One of the most challenging consequences of a natural disaster is the amount of debris it produces. It is not uncommon for one disaster to produce an amount of debris that is several times greater than the amount of annual waste produced by the affected area (Çelik *et al.*, 2015). For example, the April 25, 2015 earthquake in Nepal generated over 3.9M tons of debris in the Kathmandu Valley alone, which is the equivalent of eleven years of waste from that region (Murli Gopal Ranjitkar, 2015). The amount and composition of the debris is dependent on the built environment and the severity of the natural disaster, but typically it will be made up of construction materials (from collapsed buildings), immobilized or abandoned vehicles, and vegetation. Debris is problematic in the long term as it impedes reconstruction efforts, affects community morale, and

can pose a public risk if it contains hazardous materials (Brown, 2014). A more immediate problem is the effect of debris in the roadway on the emergency response.

In the hours and days following a natural disaster, response efforts are typically focused on performing search and rescue missions, delivering relief to those who urgently need it, transporting the injured to medical facilities, and in some cases, evacuating people if there is still an immediate threat (FEMA, 2010; Brown, 2014). Large debris piles in the roadway interfere with emergency activities, and therefore need to be either cleared to the side of the road or removed completely. To determine the location and extent of debris, a preliminary estimate is typically performed through field surveys, predictive modeling software, or an acquisition of remotely sensed data (FEMA, 2012). Field surveys are conducted by teams assigned to portions of the affected area and involve recording the location of debris and building damage, as well as estimating debris volume using measuring tapes and simple equations (FEMA, 2010). Modeling software, such as Hazards United States (HAZUS), is sometimes used to obtain a prediction of the location and extent of disaster-related damage, which can then be confirmed by deploying ground crews (FEMA, 2012). If the size of the disaster and the required information for emergency response justifies it, remotely sensed data will be collected and presented to expert analysts to extract information that is useful to emergency management officials. Remotely sensed, post-event data are often used in conjunction with pre-event imagery, critical infrastructure data, elevation data, and other useful geospatial information to map damage and debris (FEMA, 2016). Once a debris pile has been identified in an aerial or satellite image, it's volume is roughly approximated using the known pixel scale and a rudimentary volume equation, *e.g.*, for a collapsed building, volume $= \frac{\text{length} \times \text{width} \times \text{height}}{3}$ (FEMA, 2010). In some cases, to get a quick estimate of the amount of debris that needs to be removed, an estimate of the amount of debris will be obtained for a small area or length of road that represents an average amount of debris, and then the amount of debris per length or area is used to multiply by the length or area of the affected roads in an entire jurisdiction (FEMA, 2016). As one can imagine, the quantities of debris estimated using the methods mentioned above are very coarse approximations of the true volume that needs to be cleared.

There is a whole field of research devoted to developing mathematical models that determine the optimal

prioritization of roads to be serviced, such as debris clearing, road repairs, and snow removal. For example, Aksu and Ozdamar (2014) proposed a mathematical model for clearing earthquake debris from city roads in Istanbul, Turkey. The model depended on the estimated restoration time for each road, the number of machines, the length and number of work shifts, and a set of priority paths that included a debris dump site and the official evacuation route for the city. Using the model, the authors simulated an earthquake scenario for Istanbul and showed that by following the road prioritization computed by the model, that all of the roads could be cleared within three ten-hour shifts. In the case of removing debris, one of the main inputs to the model is the volume of debris in each road segment, or equivalently, the time to clear the road (Feng and Wang, 2003; Özdamar *et al.*, 2014). For these optimization models to perform correctly, they require accurate estimates of the amount of roadway debris. Additionally, manual estimation methods, such as field surveys and expert analysis, consume a lot of time and resources due to the meticulous nature of the work. In the case of providing emergency assistance, an operation being delayed due to slow debris mapping could be the difference between saving a life or not. The time and accuracy limitations of current manual mapping and volume estimation practices highlight the need for a methodology to rapidly and precisely detect and quantify debris piles in the roadway caused by natural disasters.

The increasing affordability and number of commercial providers of remotely sensed data have made the use of remote sensing for post-disaster debris analysis more feasible than ever before. Airborne platforms, such as airplanes or UAVs, can be rapidly deployed to collect data over affected regions at a relatively low cost and very fine spatial resolution (Stow *et al.*, 2015). To obtain an accurate estimate of the volume of a debris pile, height information is required. Typically, 3-D data are captured in the form of a point cloud, either directly using lidar or indirectly through reconstruction from multi-view imagery. Both methods produce point clouds that are suitable for debris detection and volume estimation, and each has its own advantages. Image-based point clouds can be much denser than lidar point clouds and retain the spectral information captured by the FPA, whereas most lidar systems only record the intensity of the backscattered laser pulse. On the other hand, reconstructing point clouds from image matching can require significant post-processing, due to the large number of overlapping images used to build the point cloud. Point clouds from lidar can

be used as soon as the data are downloaded from the sensor, which is important in a time-sensitive scenario like emergency response. Another advantage of lidar is that it is an active illumination technology, so data can be collected day or night, and illumination shadows (*i.e.*, shadows cast from tall buildings) don't create gaps in the data (Baltsavias, 1999). Regardless of how a point cloud is captured, it can be used to perform efficient post-disaster analysis of debris in the transportation network.

3.2 Literature Review

Research efforts recently have focused on developing techniques to detect and characterize debris automatically or semi-automatically from data collected immediately after the disaster. Many researchers have realized the importance of post-disaster road network assessment and have conducted studies in which they use some combination of imagery, lidar, and ancillary data to detect debris, and in some cases estimate the volume of the debris.

One of the most common approaches found in the literature to detecting roadway debris is to perform segmentation, followed by supervised, object-based classification. For example, two studies used this strategy with pan-sharpened, 0.6m QuickBird imagery collected after the Bam, Iran earthquake of 2003, and road vector data to detect debris and label roads as blocked or not. Both Samadzadegan and Zarrinpanjeh (2008) and Haghighattalab *et al.* (2010) segmented the images into homogeneous objects and then performed supervised classification with a fuzzy inference system (FIS) to label objects as different classes such as road, vegetation, and debris. Features used in the classification included both spectral metrics and Haralick texture features (Haralick, 1979). Although neither author reports a classification accuracy for debris, they both achieved over 90% accuracy in labeling roads as blocked or clear using visual inspection as their validation method.

Several other studies adopt similar approaches using imagery collected after the Haiti earthquake of 2010. Using a 0.5m resolution WorldView-2 image, Rastiveis *et al.* (2015b) trained a SVM classifier using three spectral bands, as well as six Haralick texture features on each of the bands. The output classes were masked

to the roadways using road vector data. Although a producer's accuracy of only 47% was reported for the debris class, the authors were still able to achieve a 97% classification accuracy for blockages in their test set of sixty roads. Ural *et al.* (2011) performed a massive classification of over 30 sq. km of 0.5m GeoEye-1 Haiti imagery into several classes, including debris. Using commercial object-based image analysis software, the authors used tiles of imagery and co-registered normalized digital surface models (nDSMs) to segment the area into spatially contiguous objects. Objects were manually selected for training and a hierarchical, one-class-at-a-time classification was performed using spectral, shape, texture, and height features. Debris was classified last in the hierarchy as it had no unique spectral, shape, or textural response when compared to other classes. This technique resulted in over 90% accuracy for the debris class. Although debris mapping was not restricted to the roadways, it would be trivial to crop the debris objects using road vectors. All of the studies presented thus far have required manual selection of training data and in some cases trial and error selection of segmentation parameters. While there is no doubting the utility of these methods, it is advantageous to have a fully automated method to reduce the latency between data acquisition and debris map dissemination.

Kwan and Ransberger (2010) used pre- and post-Hurricane Katrina lidar data to automatically find roads obstructed by debris through change detection. A pre-event digital surface model (DSM) was subtracted from a post-event DSM to find changes in elevation. The elevation differences were masked to the roadway using road vector data with elevation gains that spanned the width of the roadway and were between 2.5 - 5m being considered road-obstructing debris. Change detection using lidar data is an effective technique for disaster scenarios and has been widely used for building damage (Vögtle and Steinle, 2004; Rehor, 2007), but relies on the availability of pre-event data, which are not always available, particularly in developing countries. Algorithms that do not rely on multi-temporal data are therefore favorable for debris mapping immediately after a disaster.

A fully automated debris detection algorithm for very high resolution (VHR) aerial imagery is proposed in Ouzounis *et al.* (2011) and impressive results using imagery from the Haiti earthquake are presented. Images were decomposed using an efficient max- and min-tree structure that allowed them to be rapidly searched.

Debris was detected through the use of differential morphological profiles with an approximate classification accuracy of 92% using visual inspection. Using lidar data collected after the Haiti earthquake and road centerline vectors from OpenStreetMap (Ramm *et al.*, 2011), Liu *et al.* (2014) developed an application for automatically detecting blocked roadways. DSM rasters were generated from lidar data, sampled along road centerlines, to create an elevation profile through the center of each street. By comparing the elevations at the center of the street to the elevation of the road (determined by finding the minimum elevation within a buffer along the centerline), road blockages were automatically detected. The method is described as very accurate using visual inspection for validation, although it is noted that some false alarms were generated by incorrectly labeled roads in OpenStreetMap, which is an open, crowdsourced mapping effort.

There are very few studies in the literature that follow debris detection with an actual estimate of volume or characterization of individual piles of debris. Some of the studies listed above determine if roads are blocked or not, which could be used as binary input to a mathematical routing model, but can't be translated into the time it would take to clear the road. Other studies calculate a total volume of debris for a whole scene by multiplying the area of pixels classified as debris by the mean height derived from the point cloud, but these volumes are typically for whole regions, rather than individual debris piles, and are mainly used for loss estimates (Hansen *et al.*, 2007; Ural *et al.*, 2011).

Calculating the volume of tree crowns using airborne lidar is a similar problem to calculating debris pile volume. In both cases, a set of 3-D points, that typically doesn't closely resemble any shape, must be modeled in some way to calculate the volume. One technique that is useful in these types of situations is reconstruction by alpha shapes. Alpha shapes are a family of shapes that can be derived from the Delaunay Triangulation of a spatial point set (Edelsbrunner and Mücke, 1994). Many studies have used alpha shapes as a simple way to reconstruct a tree crown and estimate its volume (Yao *et al.*, 2012; Korhonen *et al.*, 2013). There are commercial software packages available that are capable of measuring volumes in point clouds. While effective, these routines require one to manually trace each object, making the process both tedious and subjective.

In this dissertation, methods found in the literature are built upon to create an unsupervised, monotem-

poral approach to roadway debris detection using airborne lidar point clouds. Point normal vectors and derived surface properties are used to identify debris points and cluster them into individual debris piles, and then alpha shapes are used to estimate the pile volume. The rest of this chapter is organized as follows: Section 3.3 introduces the study area and data used in this research, as well as a detailed description of the algorithm and workflow, Section 3.4 presents an in depth discussion of experimental results using point clouds from the Haiti earthquake of 2010, and Section 3.5 provides some conclusions from the research.

3.3 Materials and Methods

3.3.1 Debris Detection and Volume Estimation Data

The study area used to test the proposed algorithm is consists of seven sites located in the Haitian cities of Port-au-Prince and Carrefour, two regions that were heavily affected by the 7.0M_w earthquake that occurred on January 12, 2010 (see Figure 3.1). Areas with several obstructions in the roadway were identified using a map of road obstruction densities, produced by the United Nations Institute for Training and Research (UNITAR) / UINTAR Operational Satellite Applications Program (UNOSAT) from GeoEye-1 imagery acquired the day after the earthquake (UNITAR, 2010). The sites cover 0.73km^2 and were chosen because they contain a complex mixture of debris, vehicles, and people. A wide variety of roads are represented in the seven scenes, spanning everything from wide concrete roads to narrow dirt corridors. Debris piles in the roadways range from small piles of wood to large mounds of concrete that extend across the entire width of the road.

There are two data inputs to the proposed workflow: airborne lidar point clouds and road vector shapefiles. The lidar data were collected in a joint effort by Kucera International Inc. and Rochester Institute of Technology (RIT) by airplane on January 21, 2010. The lidar system used was a Leica ALS60, which operates at a wavelength of 1064nm and a pulse rate of 150kHz. The airplane was flown at an altitude of approximately 820m, resulting in an average point density of 4.2 pts/m^2 for the seven sites. Vertical point measurement accuracy is nominally 0.15m. The data were provided in the standard LAS format with X,

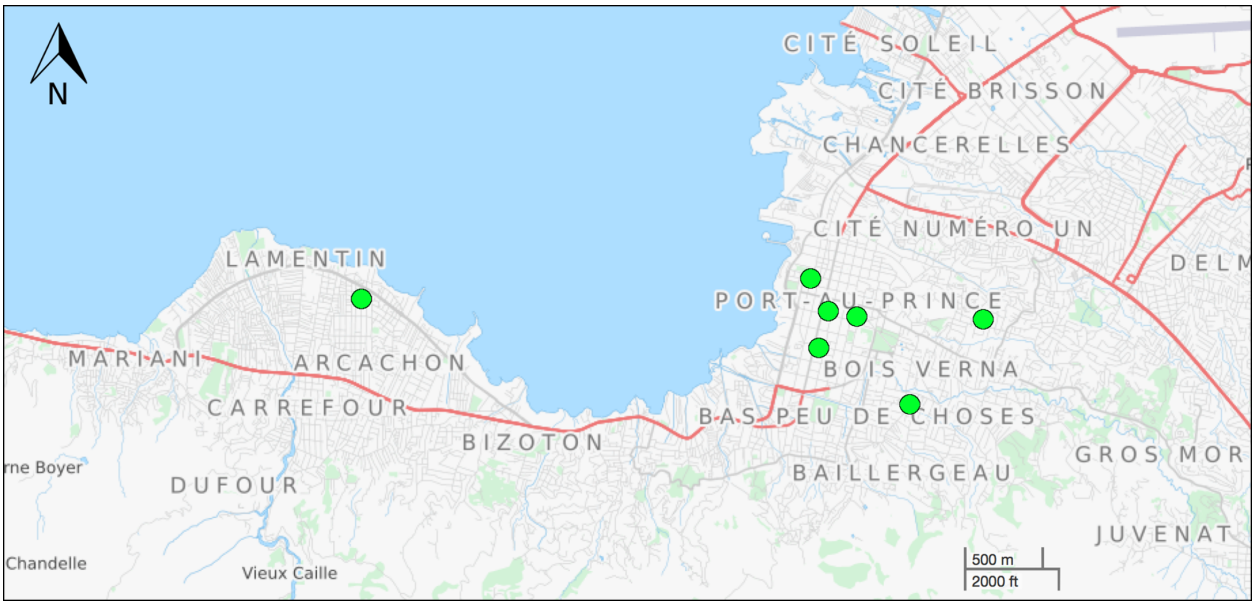
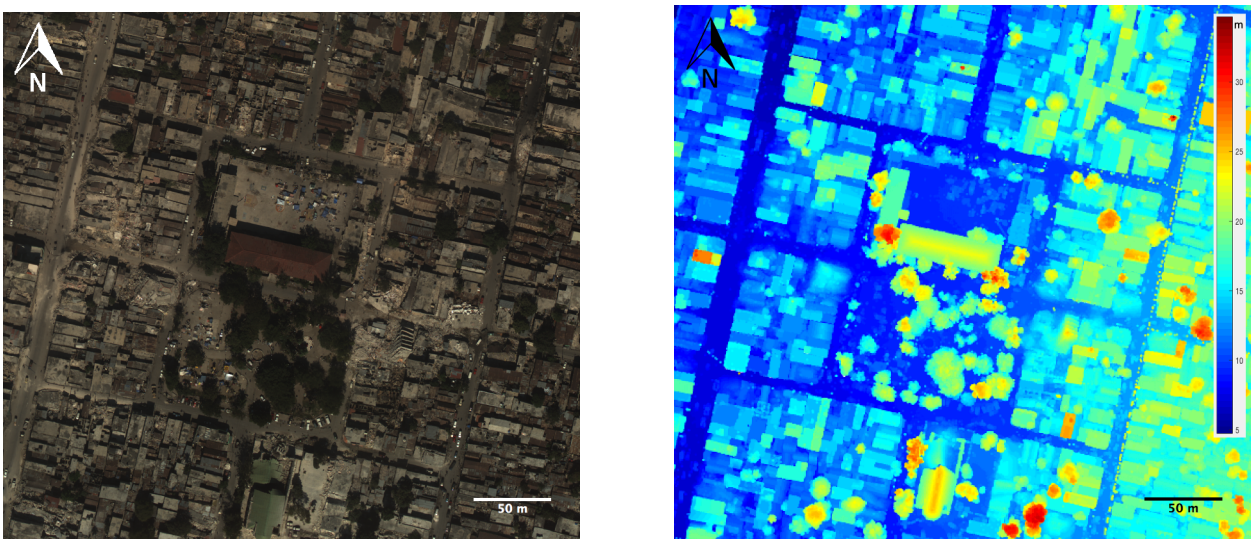


Figure 3.1: A map showing the portion of Haiti that was used in this research. The location of the seven test sites are marked by the circles.

Y, Z location, return intensity, return number, number of returns, and a vendor-supplied ground vs. non-ground classification for each point. Multispectral imagery simultaneously was collected on the aircraft by the Wildfire Airborne Sensor Program (WASP) system at a resolution of 0.15m (Messinger *et al.*, 2010). Although the imagery was not used for debris detection or volume estimation, it served as a reference for visualization and is used in figures throughout the chapter. Figure 3.2 shows both a WASP image and lidar point cloud of one of the scenes from Port-au-Prince.

The road vector shapefiles for Haiti were created by manually digitizing individual roadbeds using a combination of WASP imagery and lidar-derived intensity images. Some cities, especially in the United States, provide shapefiles of the roadbed free of charge through their local geographic information system (GIS) repository. If such shapefiles are available, they can be used directly as inputs to the algorithm.

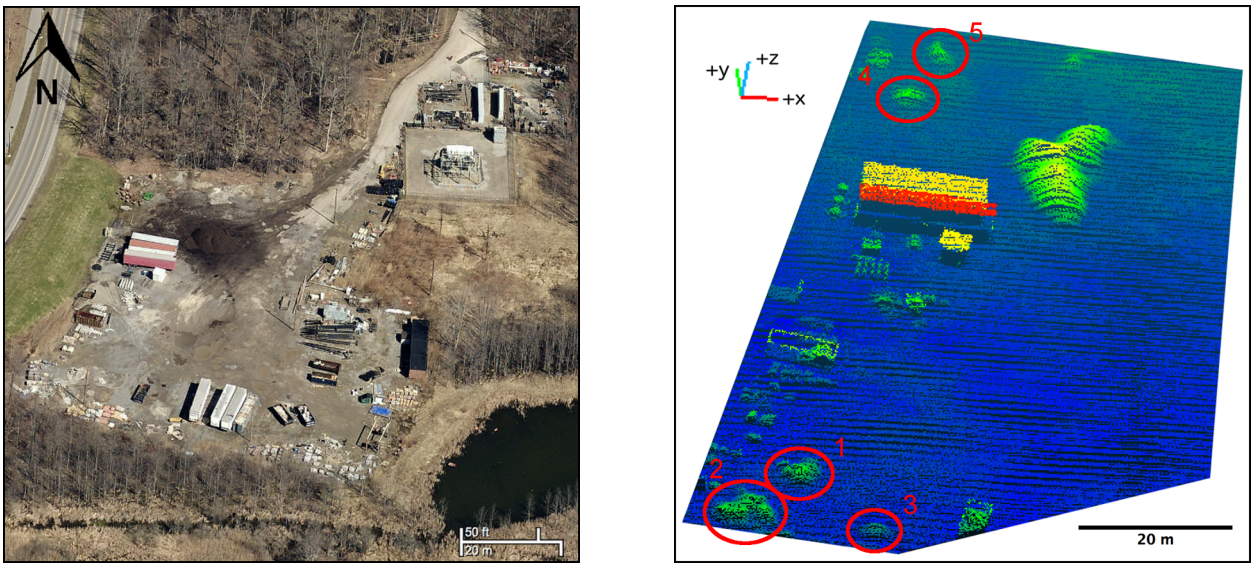
To validate the volume estimation capabilities of the proposed algorithm, an experiment was conducted in which airborne and terrestrial lidar point clouds were collected of various debris-like piles at two construction sites. By comparing volumes obtained from dense terrestrial scans (>2,000 pts/m^2 in the portion containing



(a) (b)

Figure 3.2: One of the sites from Port-au-Prince: (a) A WASP image of the scene; and (b) The corresponding lidar point cloud of the scene. The cloud is colored by height.

the piles) with those obtained from relatively dense airborne scans (between approximately 20-50 pts/m^2) the accuracy of the estimation method was analyzed. The airborne lidar system was the same Leica ALS60 system that was used for the Haiti collect. A total of 18 piles were scanned on the ground and from the air within the same 48 hours to ensure consistency. The terrestrial lidar system, which was integrated at the Rochester Institute of Technology and the University of Massachusetts, Boston, used a SICK LMS-151 laser scanner operating at 905nm (Kelbe *et al.*, 2015). The system has a maximum pulse frequency of 27kHz and a range accuracy of 30mm. The piles consisted of various materials such as concrete, gravel, and mulch. The debris piles ranged in volume from 3.04-123.11 m^3 with an average volume of 28.76 m^3, and ranged in height from approximately 1-3 m. Figure 3.3 shows an oblique airborne image and an airborne lidar point cloud of one of the sites.



(a) (b)

Figure 3.3: One of the two construction sites in Rochester, NY. The image was taken two months before the lidar data were collected, so the contents of the site differ in many spots: (a) An oblique aerial image of the scene; and (b) An airborne lidar point cloud of the scene. The five debris piles used in this study are circled in red and numbered.

3.3.2 Debris Detection and Volume Estimation Algorithm

The proposed workflow consists of three distinct sections: point cloud preprocessing, debris detection, and volume estimation. A flowchart of the method is presented in Figure 3.4.

3.3.2.1 Point Cloud Preprocessing

The inputs to the proposed workflow consist of a lidar point cloud (either raw or classified to some extent by the vendor) and a GIS file containing the road network as polygons. All points outside of the roadbed are discarded using the boundaries of the road polygons. The remaining point cloud is then filtered to remove outlier points, such as noise and power lines, using statistical outlier removal (Rusu, 2010). The statistical outlier removal stage operates by first calculating the mean Euclidean distance, $\bar{d}$, between each point, p, and its k closest neighbors. The mean (μ_k) and standard deviation (σ_k) of all of the mean distances are

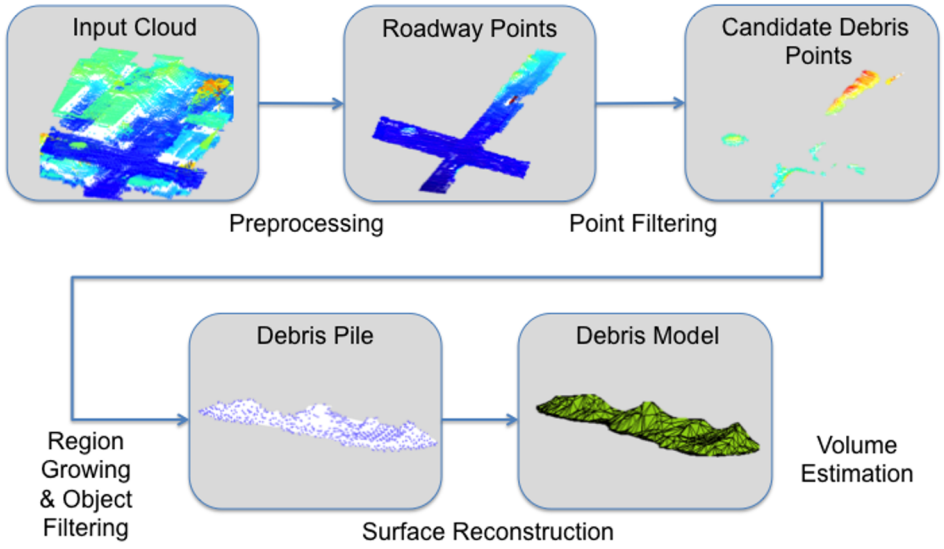


Figure 3.4: A flowchart of the debris detection and volume estimation algorithm.

calculated, and points that are considered statistical outliers are removed (equation 3.1):

$$P^* = \{p \in P | \bar{d} \leq (\mu_k + \gamma \cdot \sigma_k)\} \tag{3.1}$$

where P^* is the entire point cloud after statistical outlier removal and γ is a scalar multiplier to control the severity of point removal.

Following statistical outlier removal, a DEM is generated. Ground points and nonground points are classified using the Progressive Morphological Filter (PMF) (Zhang *et al.*, 2003). The PMF uses a morphological filter that progressively increases in size, along with elevation difference thresholds, to determine if points belong to the ground or nonground (*i.e.*, buildings, vegetation, and other objects) class. The minimum height above the ground that a point must be to be considered a nonground point is set to the vertical accuracy of the lidar system, which is typically on the order of 15 cm. Using such a low value allows low debris piles to be correctly classified as nonground. Once the ground points have been classified, a DEM is created by performing a Delaunay Triangulation with natural neighbor interpolation. The DEM point spacing is set to μ_k. The nDSM is then created by simply subtracting the point cloud from the DEM. The nDSM represents

the height of all objects above the ground, and is used later in the workflow for height thresholding. At this point the cloud is properly prepared for debris detection. Figure 3.5 shows a portion of a point cloud before and after preprocessing.

3.3.2.2 Debris Detection

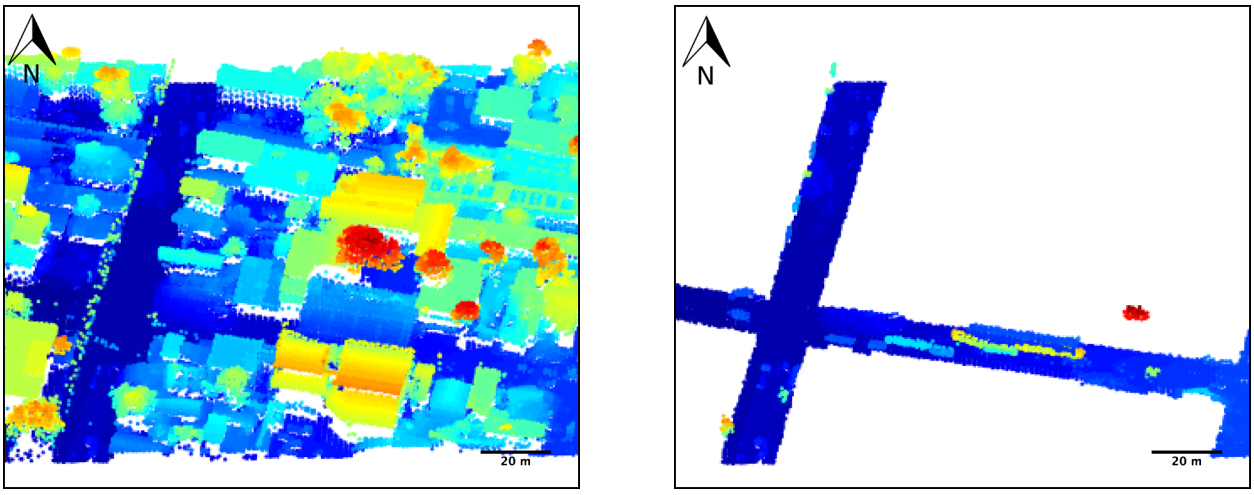
The main objectives of the debris detection stage are to classify all points as debris or non-debris and to group points that belong to the same debris pile. It is challenging to define a single set of features to identify debris, because debris objects vary remarkably in size, shape, and composition. Instead, individual points that are highly unlikely to be debris are first systematically removed using local surface properties and elevation cues. A region growing approach is then used to cluster points that belong to the same objects (*e.g.*, debris piles, cars, etc.). Finally, object-level attributes, such as height variance and number of points, are used to discard false alarms (Axel *et al.*, 2016).

One of the main benefits of working with 3-D data is the ability to estimate the normal vectors of individual points, as well as the local distribution of normal vectors. These properties give insight to the underlying geometry of the surface that each point, p, belongs to (Rusu, 2010). The normal vector of p can be calculated using all points in a local neighborhood, N_p, defined by the radius, r. For each query point, p, the neighborhood points are obtained using:

$$N_p = \{q \in P | d(p, q) < r\} \tag{3.2}$$

where q are all of the points in the entire point cloud P, and d is the Euclidean distance between two 3-D points. Eigenanalysis of the covariance matrix of N_p produces the eigenvalues $\lambda_1 < \lambda_2 < \lambda_3$. The eigenvector corresponding to λ_1 is the estimate of the point normal, $\vec{n}$. Airborne lidar data are collected from above, so the absolute value of the z-component of $\vec{n}$ is used to ensure the normal vector points outward from the surface.

The angle, θ, between each normal vector and horizontal is another useful property for point filtering:



(a) (b)

Figure 3.5: Preprocessing of a point cloud: (a) The input point cloud; and (b) The point cloud after statistical outlier removal and road point extraction.

$$\theta = 90^o - \arccos(\vec{n}_z) \tag{3.3}$$

where $\vec{n}_z$ is the z-component of the normal vector. Points sampled from a horizontal surface have $\theta = 90^o$. The neighborhood analysis is taken one step further by computing the distribution of normals in N_p. Eigenanalysis is again used, but this time on the covariance matrix of N_p, resulting in ${\lambda^n}_1 < {\lambda^n}_2 < {\lambda^n}_3$. The eigenvalue corresponding to ${\lambda^n}_2$ represents the maximum variation of normals on the Gaussian sphere (Sun and Salvaggio, 2013), and is useful for discerning points on a smooth surface, *e.g.*, road, roof, etc., from points on a rough surface, *e.g.*, a debris pile.

A majority of the non-debris points can be identified and discarded using a combination of θ, ${\lambda^n}_2$, and height above ground, h. There are two sets of points that need to be removed at this stage: low points (*e.g.*, road, sidewalk, etc.) and points that are above the ground, but on smooth surfaces (*e.g.*, trucks, road barriers, etc). Low points are removed by finding points that fall below a normal variation threshold (T_λ) and minimum height threshold (T_h) using equation 3.4. T_h is typically set to the vertical positioning accuracy

of the lidar system used to collect the data (0.15 m for this research). Smooth points are identified by using a normal variation threshold that is 5x stricter, with no height threshold (equation 3.5). In both cases, a normal angle threshold (T_θ) is included to make sure that points on highly-angled surfaces are not removed (these are much more likely to be debris):

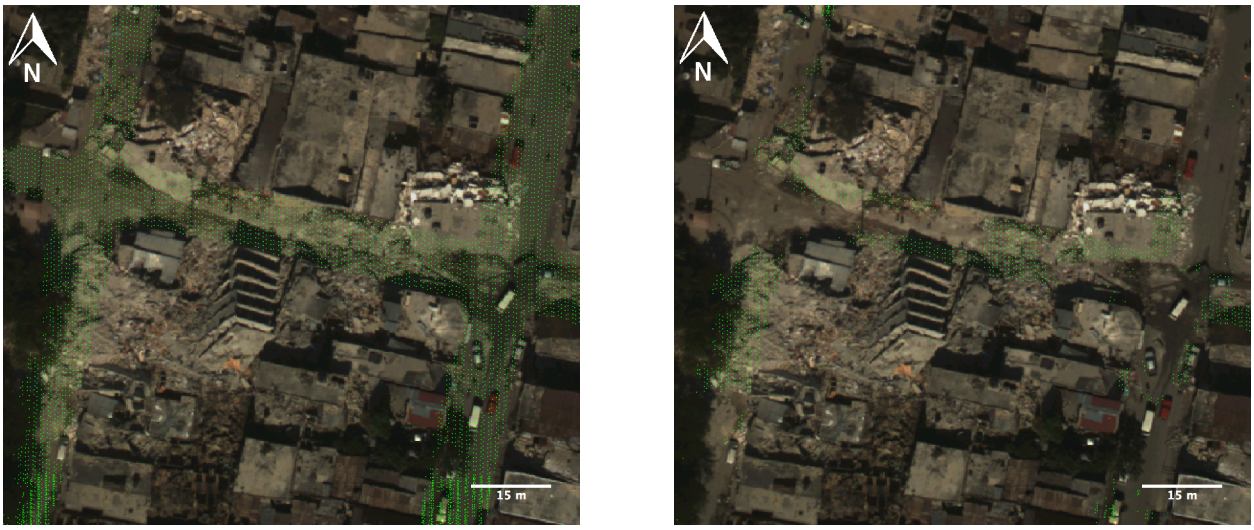
$$I_1 = \{p \in P \mid (\lambda^n{}_2 < T_\lambda) \wedge (h < T_h) \wedge (\theta > T_\theta)\} \tag{3.4}$$

$$I_2 = \{p \in P \mid (\lambda^n{}_2 < \frac{T_\lambda}{5}) \wedge (\theta > T_\theta)\} \tag{3.5}$$

where I_1 and I_2 are the indices of points classified as non-debris by equations 3.4 and 3.5, respectively. The values of T_λ and T_θ are automatically derived from critical points in the histograms of $\lambda^n{}_2$ and θ, respectively. T_h is set to the vertical point accuracy of the laser scanner.

Figure 3.6 shows an image of a debris-littered street, along with the corresponding point cloud, before and after point filtering. A majority of the points remaining after filtering belong to debris, although some points on cars and small objects near the side of the road are still sometimes present.

In order to process and analyze debris piles at the object level, individual points belonging to the same pile must be clustered together. This is accomplished through a process called *region growing*. A seed point is randomly chosen from the point cloud. The angle, β, between the normal of the seed point and the normals and each of its neighbors, which are obtained using equation 3.2, is computed. Any neighbor with a β below a certain threshold, T_β, is added to the region of the seed point. Through experimentation, a T_β of 10^o proved to be sufficient for preventing separate objects from merging. Points are iteratively added to the region until no points can be added, after which a new seed point is chosen. This process is repeated until all points have either been added to a region or discarded, because they couldn't be added to any region. A minimum region size of 15 points is used to discard regions that are too small to be significant. It should be noted that the minimum region size is only for the process of region growing. Debris piles are also discarded later on in the workflow after volume has been estimated by using a minimum volume threshold. Figure 3.7



(a) (b)

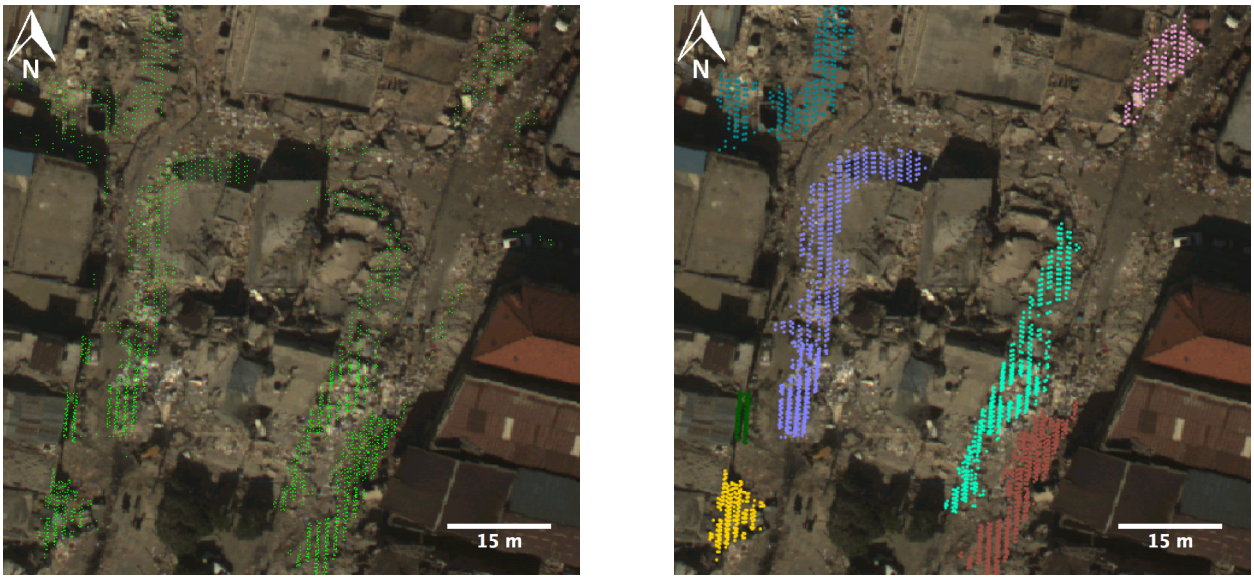
Figure 3.6: Non-debris points in a region affected by the earthquake are removed through point filtering. The points are overlaid on an aerial image for clarity: (a) Before point filtering; and (b) After point filtering.

shows the result of region growing being applied to a filtered point cloud.

At this point, the initial set of objects, *i.e.*, debris piles, have been detected. However, the results are likely to contain false alarms from small, non-debris objects such as small cars, road barriers, and medians. To filter out these objects, a metric called the height ratio (HR) is computed for each object. The height ratio is defined by Equation 3.6 and is a measure of the portion of the object that is very low to the ground. Objects that have a HR greater than a threshold, T_{HR}, are discarded as non-debris objects:

$$HR = \frac{\#P_o < 2 \cdot T_h}{\#P_o} \tag{3.6}$$

where P_o are all of the points that make up an object (produced from region growing). Figure 3.8 shows a set of objects before and after object filtering. Region growing identifies four objects in a damaged section of Haiti. Two of the objects, which are clearly false alarms, are removed by object filtering. Figures 3.8c and 3.8d show zoom views of the false alarms, which are terrain and vehicles, respectively. The objects remaining



(a) (b)

Figure 3.7: Region growing is applied to a set of filtered points in an area with several collapsed buildings. The points are overlaid on an aerial image for clarity: (a) Before region growing; and (b) After region growing.

after filtering are the set of detected debris piles. Now that the debris piles have been detected, they can be reconstructed as surfaces for volume estimation and visualization.

3.3.2.3 Volume Estimation

The last step in the workflow is to estimate the volume of each individual debris pile. An assumption is made that the bottom of the debris pile sits flush on the road beneath it. To this end, the DEM points that fall within the boundaries of the 2-D projection of a debris pile are appended to the pile for improved volume estimation. Each pile is then modeled as a 3-D surface using alpha shapes (Axel *et al.*, 2016). Alpha shapes are parameterized by an alpha radius, α, that controls how tightly the surface wraps around the set of points. For volume estimation, it is important that the alpha radius is large enough to form a closed surface around the points. Setting α to $3\mu_k$ is sufficient for generating closed surfaces around the debris

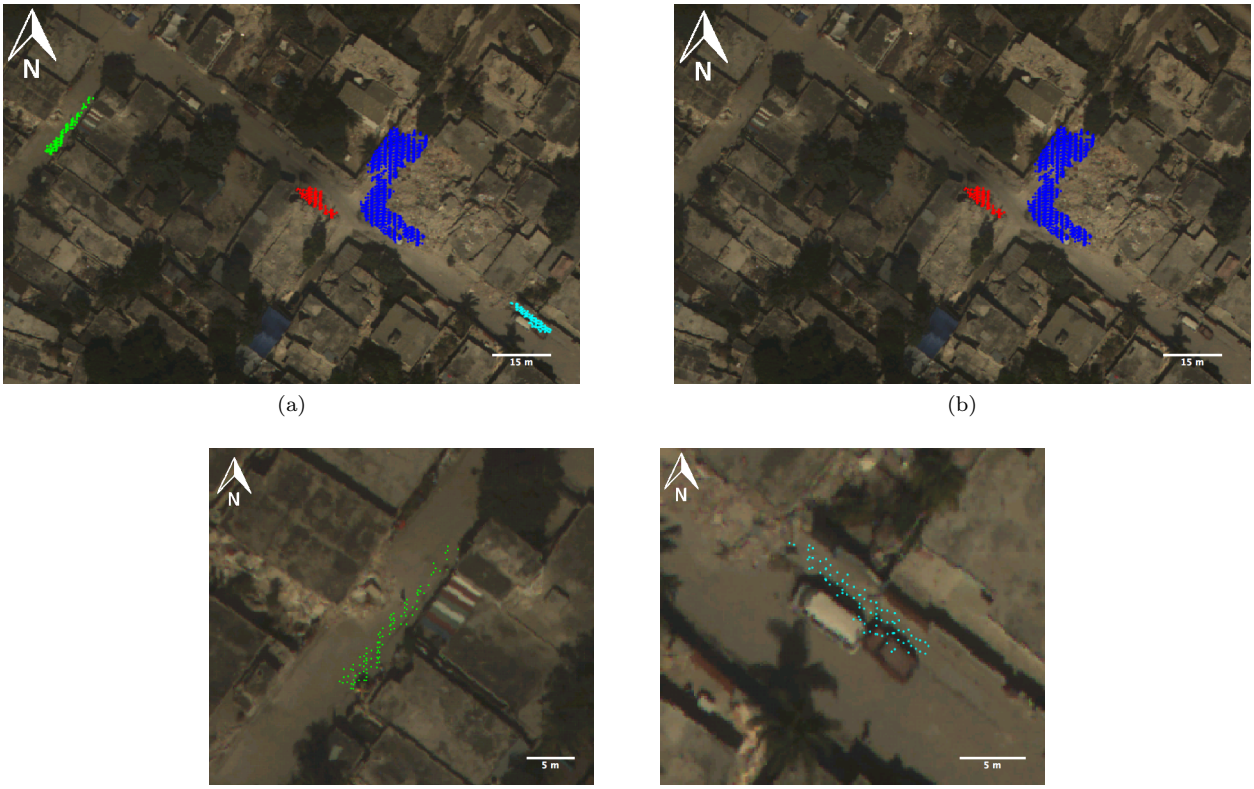


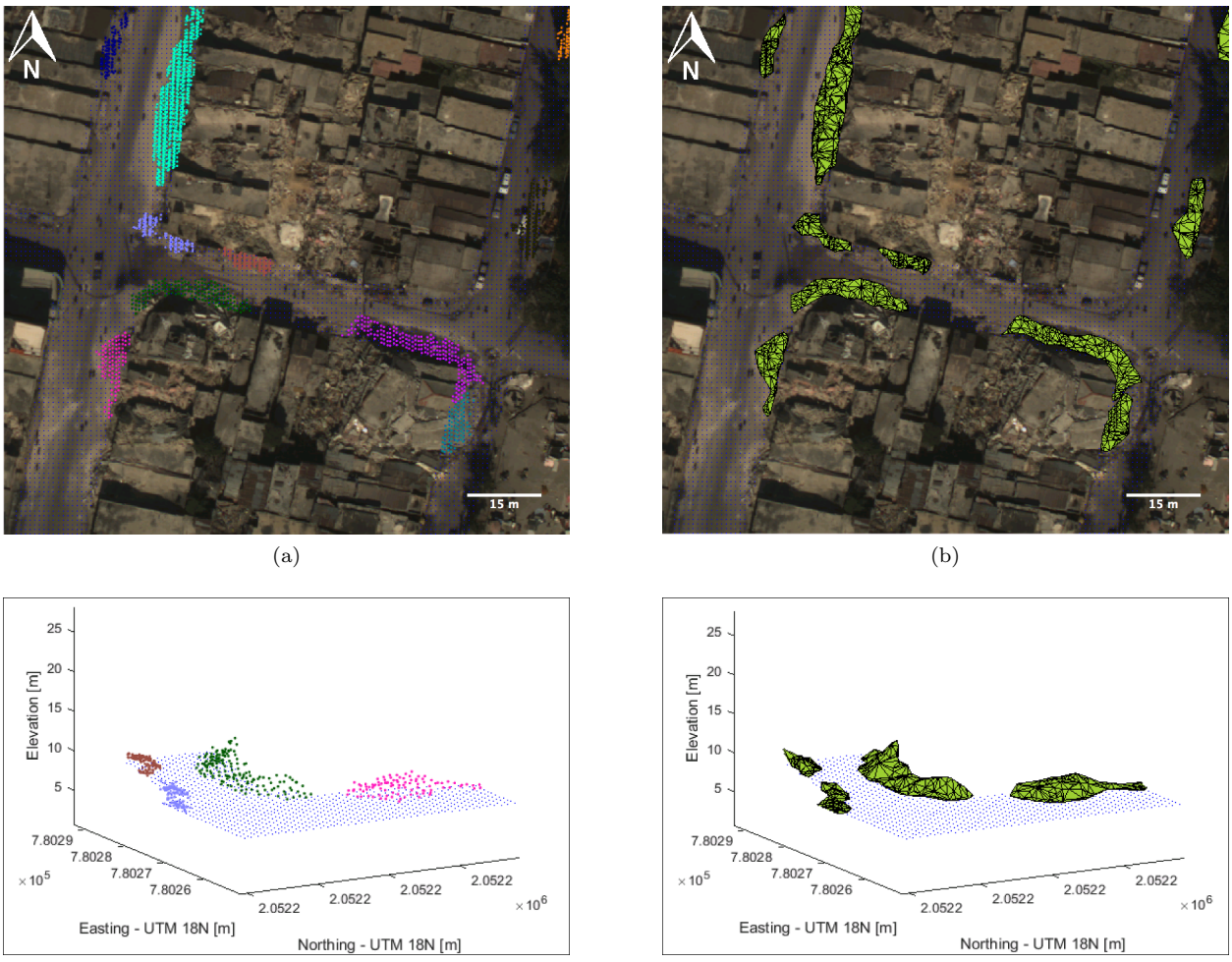
Figure 3.8: Object filtering is performed and removes two non-debris objects that were obtained from region growing: (a) Four objects are created during region growing; (b) Two of the objects are removed as a result of object filtering; (c) A zoom view of the first removed object shows that it was a section of terrain that was mistaken for debris; and (d) A zoom view of the second removed object reveals that it was two vehicles clustered together.

piles. Once a debris pile has been modeled using alpha shapes, it can be treated as a set of tetrahedrons whose individual volumes are computed and summed to find the total volume for the pile. This process is carried out for each debris pile. Figure 3.9 shows a set of roadway debris piles in Haiti before and after alpha shape reconstruction. At this point the debris piles and volumes can be exported as a GIS file (e.g. shapefile) and opened as a debris map in GIS software.

3.3.3 Experimental Validation

3.3.3.1 Validation of Debris Detection

Reference data to validate the debris detection results were created by manually tracing every pile visible in the high resolution (0.15m) WASP imagery. The bounding polygons of the 2-D projections of the detected debris piles were intersected with the set of validation polygons for each scene. If a debris pile polygon detected by the algorithm overlapped with a validation polygon, then it was classified as a true positive (TP). If there was no overlap between the detected debris polygon and any of the validation polygons, then it was labeled as a false positive (FP). If a validation polygon did not overlap any of the detected debris polygons, then it was classified as a false negative (FN). The validation polygons were sometimes larger than the detected polygons because debris that was present in the nadir aerial imagery was very low to the ground and discarded during the point filtering stage of debris detection. The low points were intentionally discarded as insignificant contributions to the overall debris pile. Therefore, no threshold was placed on the amount of overlap between the two sets of polygons to achieve a correct detection. During validation, if a digitized pile was found to be completely below the minimum height threshold, it was deleted and the results were recalculated. The TP, FP, and FN counts were used to characterize the performance in terms of completeness, correctness, and quality. *Completeness* reflects the percentage of validation debris piles that were detected by the algorithm (equation 3.7). *Correctness* reflects the percentage of detections that were true debris piles (equation 3.8). *Quality* is a measure of overall performance that takes into account both the completeness and the correctness of the results (equation 3.9) (Heipke *et al.*, 1997).



(a) (b)

(c) (d)

Figure 3.9: A set of roadway debris piles in Haiti as reconstructed using 3-D alpha shapes. The piles are overlaid on an aerial image of the scene for clarity. The road points are shown in blue: (a) Each set of colored points represents an individual debris pile detected by the proposed workflow; (b) The debris piles reconstructed as alpha shapes; (c) A perspective view of the debris piles as points. Note that the view is rotated from (a), but the colors correspond to the same piles; and (d) A perspective view of the debris piles as alpha shapes.

$$Completeness = \frac{TP}{TP + FN} \times 100, \tag{3.7}$$

$$Correctness = \frac{TP}{TP + FP} \times 100, \tag{3.8}$$

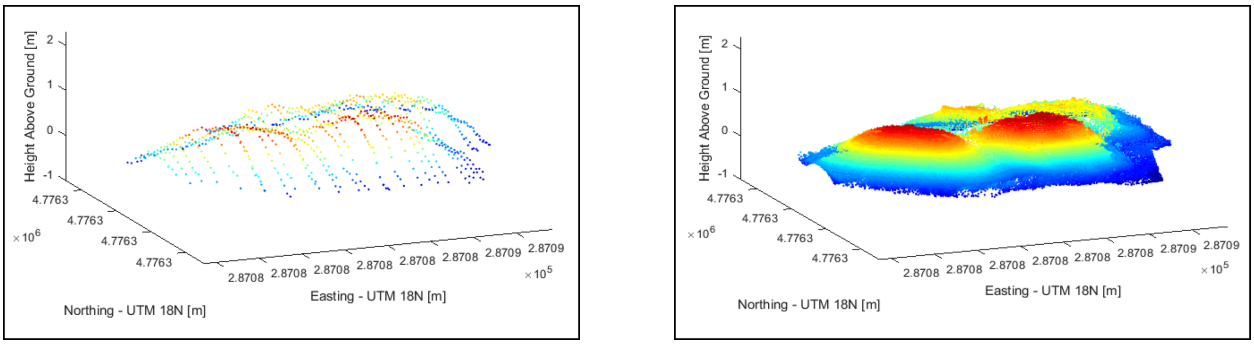
$$Quality = \frac{TP}{TP + FP + FN} \times 100 \tag{3.9}$$

3.3.3.2 Validation of Volume Estimation

There was no direct manner in which to validate the volumes estimated for the debris piles in Haiti, because no field data were ever published for the volume of individual piles. Even if individual pile volume data were published, it would have to have been measured on the same day as the lidar data were collected to be consistent. Rather than validating the volumes directly, an experiment was conducted to measure the agreement between volumes obtained from high resolution, ground-based scans and airborne scans of debris piles from construction sites. Although this method did not guarantee the accuracy of the volumes, it characterized the volume estimates from airborne lidar data in comparison to what can be obtained on the ground. Terrestrial and airborne laser scans were collected of debris piles at two locations in Rochester, NY. The two sets of scans were preprocessed to a state in which each pile could be reconstructed using alpha shapes and an estimate of volume could be obtained. Figure 3.10 shows both airborne and terrestrial point clouds of the same debris pile to highlight the difference in point densities of the two modalities.

The airborne lidar data did not require preprocessing before being used as an input to the proposed workflow. However, in order to locate the debris piles from the terrestrial lidar data into the same coordinate system, it was beneficial to segment each of the airborne piles into its own point cloud. Each of the debris piles in the airborne point clouds were extracted by simply delineating a polygon around the pile and extracting points within the polygon boundaries. The piles were easily distinguished from the flat parking lots on which they were located by coloring the point cloud by height.

To capture the entirety of each debris pile with the terrestrial lidar, between three and nine scans had to be collected, depending on the size of the pile. The individual scans were registered into a composite scan



(a) (b)

Figure 3.10: Point clouds of a debris pile from a construction site in Rochester, NY. The terrestrial cloud is much denser with 59,348 points compared to the 899 points of the airborne cloud: (a) Airborne lidar point cloud; and (b) Terrestrial lidar point cloud.

by manually selecting matching points between scans, which was trivial due to the placement of reflective markers in the scene. An image of one of the gravel piles, along with the corresponding terrestrial scan, can be seen in Figure 3.11.

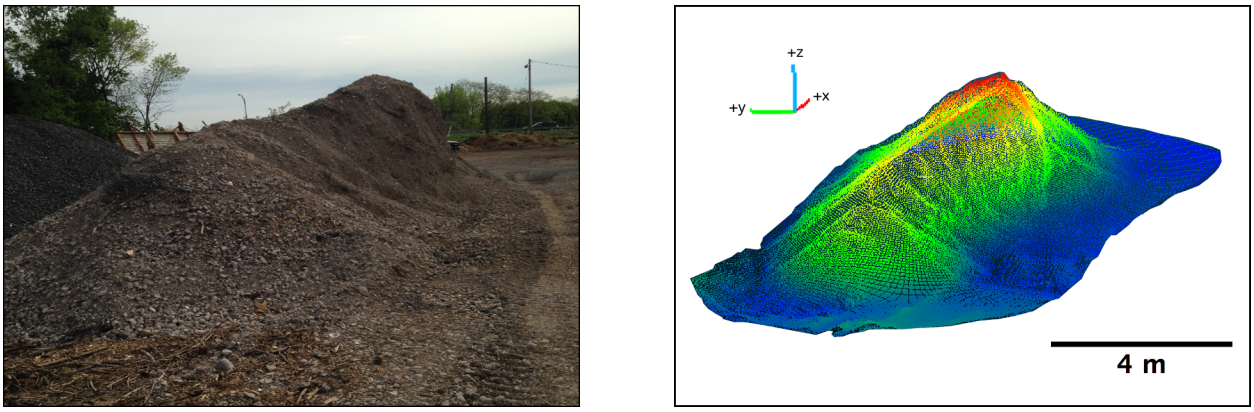
The coordinate system of the terrestrial scans is arbitrary and relative to the scanner. For ease of comparison and visualization, the scans were transformed into the same coordinate system as the airborne scans, which is Universal Transverse Mercator (UTM) Zone 18 North. This was accomplished for each pile by:

1. Rotating the terrestrial lidar scan so that positive x and positive y matched north and east, respectively.

2. Finding the highest point in the terrestrial scan and subtracting the respective x- and y-coordinates from the entire pile, so that its x- and y-coordinates at the highest point became $(0, 0)$.

3. Finding the highest point in the airborne scan and adding its x- and y-coordinates to the entire terrestrial scan.

4. Subtracting the height of the highest point in the terrestrial scan from the entire pile, so that the maximum height was zero.



(a) (b)

Figure 3.11: A large pile of gravel at one of the construction sites: (a) An image of the gravel pile; and (b) A point cloud of the gravel pile created by registering several terrestrial lidar scans.

5. Adding the height of the highest point in the airborne scan to the entire terrestrial pile so that the highest points, and presumably the rest of the pile, are at the same elevation as the airborne scan.

6. The bounding polygon of the corresponding airborne scan was used to bound the terrestrial scan, so that exactly the same area is captured in both scans.

To enclose the piles with a bounding surface on the bottom, a plane was fit to the bottom of the pile using least squares fitting. This was accomplished by projecting the points into 2-D and taking the boundary points. All of the debris piles scanned in this experiment were largest at the bottom, so these points were always at the minimum height of the pile. Points were sampled in a 0.25m grid on the bottom plane and appended to the pile.

At this point, both sets of debris pile scans were ready for volume estimation. Each of the two airborne point clouds (one for each construction site) in their entirety were input as is to the debris detection and volume estimation algorithm. The debris piles were located in parking lots, rather than roads, so no road vector input was used in this case. The algorithm produced a set of debris pile point clouds, as well as a set of alpha shape reconstructions of the piles. The algorithm detected many more debris piles than were actually scanned, since only a portion of the debris piles at each site were scanned due to time constraints,

and in some cases, piles that were in close proximity to one another were segmented as one large pile. To enable accurate comparison of the volume estimates, only the piles matching the terrestrial piles could be considered. To ensure that this was the case, the piles were segmented using the bounding polygons of the terrestrial piles and then reconstructed again as alpha shapes. Once the piles were reconstructed as alpha shapes, the volume was computed by simply summing the volumes of the individual tetrahedrons that make up the alpha shape. The debris piles from the terrestrial lidar scans were already segmented into individual piles, so they could not be used as inputs to the algorithm. Instead, they were simply reconstructed as alpha shapes and the volumes were obtained.

3.4 Results and Discussion

The proposed debris detection and volume estimation workflow was tested using manually digitized road polygons and airborne lidar data for seven scenes in Port-au-Prince and Carrefour, collected after the Haiti earthquake in 2010. The scenes were chosen because they all contained multiple roadway debris piles of varying shapes and sizes.

3.4.1 Debris Detection Results

The debris detection results for all seven of the scenes are displayed in Table 3.1. Figure 3.12 shows a comparison of the validation polygons and the bounding polygons of the debris piles detected by the algorithm for two of the Port-au-Prince sites. The proposed algorithm achieved an overall quality score of approximately 88% when tested on seven scenes (containing a total of 88 validation debris piles), indicating that it is capable of satisfactory performance in automated roadway debris pile detection. Even more importantly, the algorithm achieved a completeness score of over 98%. The completeness metric penalizes false negatives, which in the case of debris detection for emergency response, are arguably more costly than false positives. A false negative could result in an ambulance that is transporting critically injured survivors, to incorrectly choose a route to the hospital that is blocked by debris. Of the 88 validation debris piles, only one was

omitted. The debris pile that was omitted was a small portion of a roof that broke off and landed in the roadway. The portion of the roof remained almost completely planar, and was filtered out due to its smoothness and normal vector angle. Although such a clean break is rarely seen in the data, it does highlight one failure case of the algorithm.

Eleven false positives were detected across the seven test scenes, resulting in a correctness score of about 89%. All of the false positives resulted from large individual vehicles or multiple vehicles parked very close to each other. At the point density of the Haiti data, the vehicles are often undersampled in such a way that they closely resemble debris piles. This is especially true if there are multiple vehicles parked very close together. In these cases the vehicles tended to get clustered together in the region growing process and subsequently were mistaken for debris. Two examples of clusters of cars being mistaken for debris can be seen in Figure 3.13, as well as a side profile of the lidar points showing the lack of separation between the vehicles. Although only eleven false positives were produced compared to the hundreds of vehicles present in the test data, they could still pose a problem if time was wasted sending a cleanup crew to a road that didn't actually contain debris. However, the false positive debris piles caused by vehicles tended to be much smaller than the actual debris piles that resulted from collapsed buildings. To reduce the number of false alarms, one could simply raise the minimum volume threshold when running the algorithm. Additionally, if lidar data are collected at a higher point density, it is far less likely that the vehicles would be clustered

Table 3.1: Debris detection results for all seven of the scenes.

Site #	**# Validation Piles**	**TP**	**FP**	**FN**	**Completeness [%]**	**Correctness [%]**	**Quality [%]**
1	21	24	3	0	100.00	88.89	88.89
2	5	5	0	0	100.00	100.00	100.00
3	14	13	0	0	100.00	100.00	100.00
4	9	8	1	0	100.00	88.89	88.89
5	11	11	3	0	100.00	78.57	78.57
6	7	7	1	0	100.00	87.50	87.50
7	21	24	3	1	96.00	88.89	85.71
Total	88	92	11	1	**98.92**	**89.32**	**88.46**

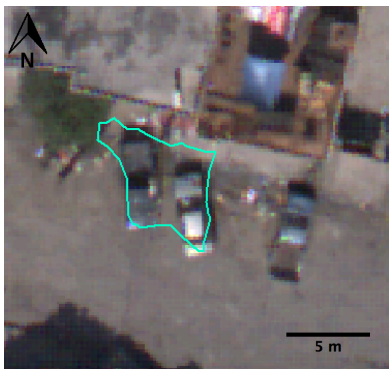
(a) (b)

Figure 3.12: Validation debris pile polygons (red) and polygons of the debris piles detected by the proposed algorithm (teal) are overlaid on aerial imagery: (a) Results for site six; and (b) Results for site one.

together and mistaken for debris. The introduction of imagery (multispectral or even color) could also help to reduce false positives.

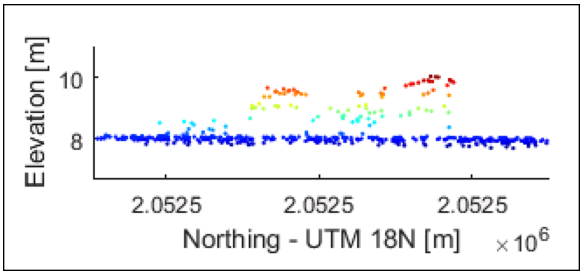
3.4.2 Volume Estimation Results

A comparison of the volumes computed for the airborne and terrestrial scans is presented in Table 3.2. At full density, *i.e.*, 52.3 pts/m^2 and 20.6pts/m^2 for lots 1 and 2, respectively, the algorithm calculated volumes with a mean error of 1.25 $\pm$ 0.9m^3 and a maximum error of 3.33m^3. The absolute value of the volume differences was used to obtain the mean error. This translates into a mean error of 9.83 $\pm$ 8.58%. Debris-clearing rates of typical machinery are in excess of 200 m^3/hr (Feng and Wang, 2003), so an estimation error of $\approx$ 3m^3 would likely be insignificant to a mathematical model used for prioritization of response activities. If the analysis were limited to debris piles with a ground-measured volume of 20m^3 or higher (this is the minimum volume threshold used for the detection testing), the mean volume error becomes 1.71 $\pm$ 1.06m^3, translating to a mean error of 4.26 $\pm$ 3.47% (see Table 3.3). These results suggest that at these resolutions, airborne



(a)

(b)



(c)

Figure 3.13: Cars parked near each other and undersampled are the main source of false positives in the debris detection results for the test scenes. The detected debris pile boundaries are shown in teal: (a) Two cars parked near each other in scene seven are mistaken for a debris pile; (b) Two cars parked extremely close to each other in scene six are mistaken for a debris pile; and (c) The lidar points from (b), which show the undersampling of the vehicles that were mistaken for debris.

lidar scans can be used as inputs to the proposed algorithm and obtain reasonably accurate estimates.

Although the results presented in Table 3.2 suggest that the proposed algorithm is able to produce accurate estimates of debris pile volumes at full resolution (52.3 pts/m^2 and 20.6pts/m^2), it is necessary to investigate the accuracy of estimates at the resolution of the Haiti point clouds (4.2 pts/m^2, on average). To accomplish this, the point clouds were randomly subsampled from 20.6 pts/m^2 to 2 pts/m^2 in increments of 2 pts/m^2 and the same volume comparisons were performed. Only debris piles with a ground-measured volume of 20m^3 or greater were considered for this analysis.

Figure 3.14 show plots of the mean volume difference in terms of m^3 and percentage. As expected, the mean error decreases as a function of point density. At the mean point density of the Haiti data (4.2 pts/m^2) the error is about 5.75m^3, or 14%. The highest error of $\approx$ 9m^3, or 21%, was obtained at the lowest tested density of 2 pts/m^2. The lowest error of $\approx$ 1.5m^3, or 5%, was measured at 18 pts/m^2. Although the lowest error should have been recorded at the highest density, the randomness of the subsampling may have contributed to a slightly lower error at 18 pts/m^2, rather than 20.6 pts/m^2. Debris piles tend to have a mixture of both small and large construction materials, which form arbitrary shapes that cause cavities and peaks. The lower density point clouds undersample these unique features and they are either missed altogether or poorly represented in the alpha shape reconstruction, resulting in a small over- or underestimation of the true volume. Figure 3.15 demonstrates the effect of undersampling the debris pile using an example from one of the construction sites in Rochester, NY. Comparing Figures 3.15a-3.15b, as well as Figures 3.15c-3.15d, it is clear that several features were misrepresented or completely absent in the downsampled point cloud, and thus resulting alpha shape. However, despite the point reduction of over 300x from the high to low density point clouds, the estimated volumes were only different by 3.2m^3.

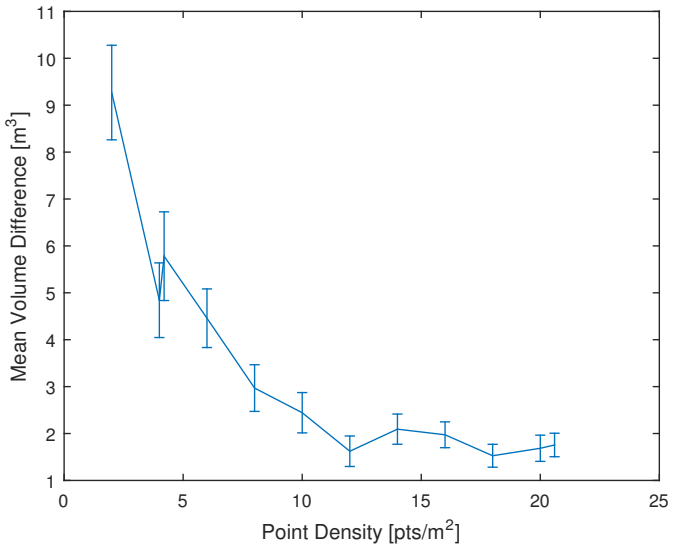
The results show that even at low point densities, volumes can be estimated at an accuracy that makes them useful such that both a human or a mathematical model could use them to make important emergency response decisions. Regardless, increased accuracy is always desired. The simplest method for increasing volume estimation accuracy is to simply collect lidar data at a higher point density. A higher point density would benefit not only volume estimation, but debris detection as well. The improved sampling that high

Table 3.2: A comparison of volume estimates for debris pile point clouds obtained using terrestrial lidar and airborne lidar.

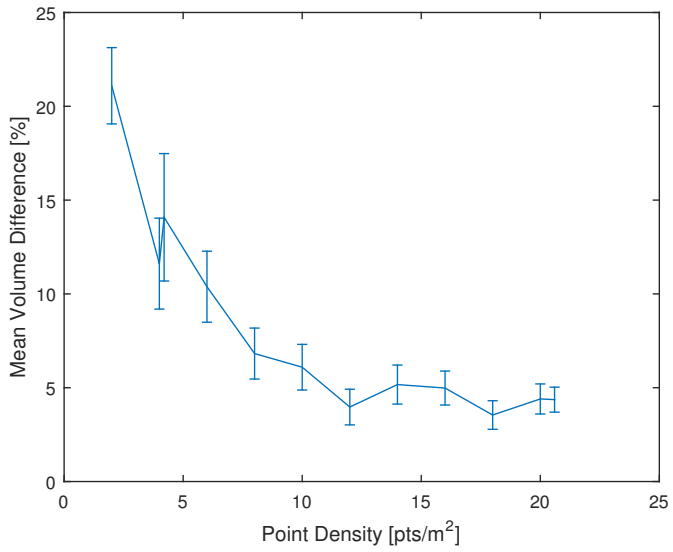
Pile #	**# Points (Ground)**	**# Points (Air)**	**Volume (Ground) [m^3]**	**Volume (Air) [m^3]**	**Volume Difference [m^3]**	**Volume Difference [%]**
1	37,367	529	10.26	10.78	-0.52	5.03
2	129,688	1,132	22.62	22.82	-0.20	0.88
3	74,199	313	4.13	3.41	0.71	17.33
4	87,192	398	5.09	4.74	0.35	6.84
5	43,891	303	3.04	2.79	0.25	8.21
6	107,566	1,210	75.81	77.78	-1.97	2.60
7	96,333	1,813	100.19	100.58	-0.39	0.39
8	110,248	1,088	47.23	50.00	-2.77	5.86
9	37,382	188	5.82	3.95	1.87	32.07
10	59,348	899	25.44	24.00	1.44	5.68
11	53,419	346	12.00	11.36	0.65	5.38
12	47,209	787	32.45	29.12	3.33	10.27
13	56,451	202	5.22	4.46	0.76	14.59
14	60,483	814	25.25	26.99	-1.75	6.92
15	35,678	224	6.27	5.44	0.83	13.22
16	182,391	2,168	123.11	121.24	1.86	1.51
17	33,535	168	7.06	5.23	1.83	25.89
18	17,205	210	6.71	5.74	0.96	14.34
				Mean	1.25	9.83
				St. Dev.	0.90	8.58

Table 3.3: A comparison of volume estimates for debris pile point clouds obtained using terrestrial lidar and airborne lidar. Only debris piles with a ground-measured volume of 20 m^3 or greater are included in this table.

Pile #	**# Points (Ground)**	**# Points (Air)**	**Volume (Ground) [m^3]**	**Volume (Air) [m^3]**	**Volume Difference [m^3]**	**Volume Difference [%]**
1	129,688	1,132	22.62	22.82	-0.20	0.88
6	107,566	1,210	75.81	77.78	-1.97	2.60
7	96,333	1,813	100.19	100.58	-0.39	0.39
8	110,248	1,088	47.23	50.00	-2.77	5.86
10	59,348	899	25.44	24.00	1.44	5.68
12	47,209	787	32.45	29.12	3.33	10.27
14	60,483	814	25.25	26.99	-1.75	6.92
16	182,391	2,168	123.11	121.24	1.86	1.51
				Mean	1.71	4.26
				St. Dev.	1.06	3.47



(a)



(b)

Figure 3.14: Plots of volume estimation error as a function of point density in terms of m^3 and percentage: (a) Volume error expressed in m^3; and (b) Volume error expressed in %.

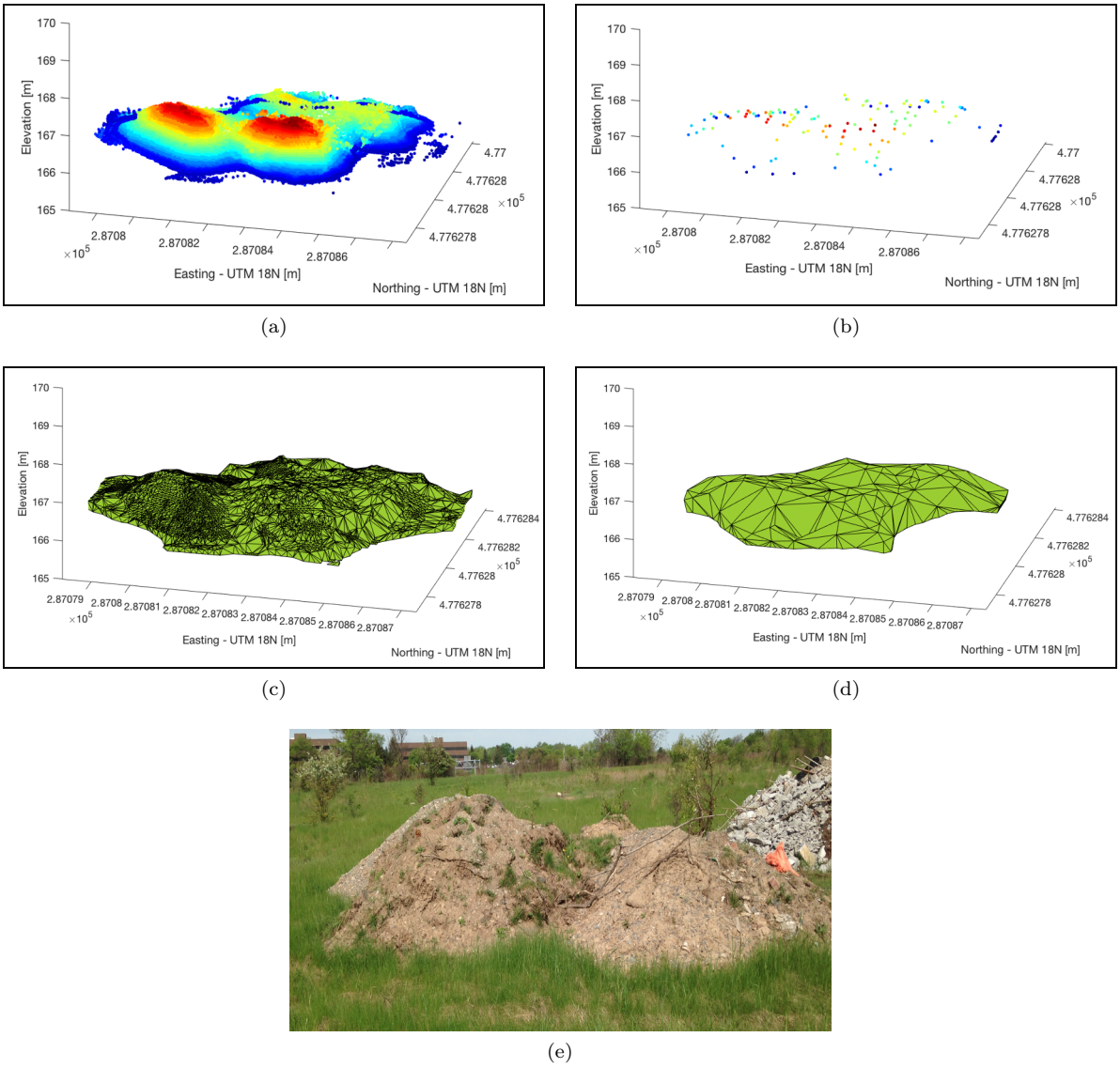


Figure 3.15: A comparison of high and low point density debris points and alpha shapes using a pile from the construction site in Rochester, NY: (a) The terrestrial lidar scan of the pile contains 41,179 points; (b) The airborne lidar scan of the pile, downsampled to the density of the Haiti data, contains 128 points; (c) The alpha shape constructed from the high density data; (d) The alpha shape constructed from the low density data; and (e) An image of the debris pile.

point density achieves, helps to separate false positive objects, such as vehicles, from true debris piles. Alternatively, a more complicated alpha shape reconstruction, using an alpha radius adaptive to the local point density, could provide a more accurate representation of the underlying surface and therefore a better volume estimation.

3.5 Conclusions

This research assessed a novel roadway debris detection and volume estimation workflow, developed to enable the use of airborne lidar data for rapid transportation network analysis following a natural disaster. Local surface features and region growing were utilized to extract roadway debris piles and then the piles were reconstructed as surfaces using alpha shapes, allowing for accurate volume estimation. The proposed technique allowed for successful extraction of over 98% of the debris piles present in a dataset collected after the major 2010 earthquake in Haiti. A correctness of 89.32% and quality of 88.46% were obtained when the algorithm was evaluated on a set of 88 debris piles in lidar point clouds from seven scenes. False positives from vehicles were present in some cases when the vehicles were undersampled due to low point density. An additional experiment was conducted to investigate the effect of point density on the accuracy of volume estimation. The point density experiment showed that at the low density of the Haiti lidar data, volumes were estimated with average errors of $\approx 6\text{m}^3$ or 14%, and significantly lower errors as point density increased. These results demonstrate the feasibility of a fully automated approach to roadway debris detection and volume estimation, which could have significant implications to the manual, labor-intensive methods currently in use. The outputs from the proposed algorithm could easily be used to create GIS-ready maps to disseminate emergency management officials (see Appendix A), or feed directly into mathematical optimization routines for routing and debris clearing (see Appendix B).

Chapter 4

Detection and Depth Estimation of Roadway Flooding using Airborne Imagery and Digital Elevation Models

4.1 Introduction

Floods are amongst the most impactful natural disasters, affecting more people globally each year than any other natural disaster (RedCross, 2012). River floods occur when natural or artificial barriers are overtopped by heavy rains, spreading water over the flood plain and endangering nearby communities. Built up areas in close proximity to the flooding can increase runoff by as much as six times of what would occur naturally, leading to loss of life and immense amounts of damage. As a result, flooding is one of the most studied weather releated phenomena. Entire fields of research are devoted to flood preparedness and mitigation through precipitation monitoring and forecasting, hydrologic modeling, structural reinforcements, and disaster planning (Jeyaseelan, 2003). Despite these preventative measures, floods still affect tens of

millions of people globally every year (RedCross, 2012). Once a flood has struck, it is vital to monitor and map the situation in real time in order to locate and identify affected regions and initiate the required response (*i.e.*, diversion of water, evacuation, rescue).

Conventional hydrological monitoring systems are of limited use in flood mapping and emergency response. Monitoring and stream gauging stations provide useful historical data, but the sparse data they provide lack the spatial flood characterization needed to provide a comprehensive overview of the situation. The gauging stations give information about the water height at their location, but no information about the extent of the flood. Additionally, a wide range of local factors, such as slope of terrain, drainage network, and land cover, need to be taken into account to make sense of the measurements produced by the stations (Klemas, 2014). Remote sensing platforms, such as satellites and aircraft, can be used in place of (or in conjunction with) hydrological monitoring systems to provide timely information to emergency management personnel. The most important data are the extent of flood inundation and the depth of the water, specifically in the roadways. The road network is the lifeline in the case of a flood, it dictates safe routes for evacuation, delivery of aid, and planning of search and rescue missions.

Both microwave and optical remote sensing have been established by many studies as effective means of flood mapping (Klemas, 2014; Musa *et al.*, 2015). Active microwave sensing modalities, such as Synthetic Aperture Radar (SAR), emit radar signal towards a surface and record the reflected energy. Strong radar absorption by water, as well as the ability to operate at any time of day and penetrate clouds, make SAR a popular choice for flood extent mapping. Various studies have effectively leveraged satellite SAR imagery for creating flood maps in the past two decades. Some examples include: segmentation thresholding of SAR images (Long *et al.*, 2014), change detection with multi-temporal interferometer SAR images (Dellepiane *et al.*, 2004), and neural networks (Skakun, 2012). As with all imaging modalities, microwave sensing is not without limitations. Surrounding or emergent vegetation, wind, and turbulence can all cause significant backscatter, thus affecting the accuracy of the radar measurements (Smith, 1997). Quality of SAR data can also be limited by wind and rain, incidence angle, and the polarization mode using for acquisition (Musa *et al.*, 2015).

Optical sensors rely on illumination from the sun to be reflected off of a surface in order to image them in visible and infrared spectral bands. Sensors capable of recording NIR energy are ideal for flood mapping due to strong absorption in water and reflection by land and vegetation (Klemas, 2014). Airborne and both medium and high resolution satellite sensors have been used for mapping various floods that have occurred across the globe in the past few decades. Examples of flood mapping with optical imagery include: change detection using combinations of spectral bands in two Landsat 7 images (Wang *et al.*, 2002), detection of flood regions by thresholding a Difference of Normalized Difference Water Indices (DNDWI) image calculated from Landsat 5 imagery (Ogashawara *et al.*, 2013), and a signature-matched detection approach using a manually selected flood pixel in an flood image acquired from an airborne sensor (Sun *et al.*, 2012). One downfall of optical sensors on satellites is that they cannot observe through clouds, which are often present due to the storm systems that lead to floods. Although airborne platforms can get around the cloud problem, all optical systems can only image during daylight and they cannot penetrate obscurations, such as vegetation.

Both microwave and optical remote sensors are valid modalities for flood mapping and depth estimation. The research presented here focuses only on airborne optical sensing, because it provides the high resolution and flexible control over all factors in image acquisition that are beneficial in mapping and estimating depth of flooded roadways during emergency response. The next section presents a detailed literature review of flood mapping and water depth estimation using overhead optical imagery.

4.2 Literature Review

4.2.1 Flood Extent Mapping

A wide variety of flood mapping algorithms using overhead optical imagery exist in the literature. Although there are major differences in terms of ground sample distance (GSD) between satellite, airborne, and UAV mapping approaches, all of the algorithms can be separated into two classification categories: supervised and unsupervised.

Supervised methods for flood mapping use manually labeled training data to train a classifier. For exam-

ple, Amini (2010) used pansharpened, 1m resolution color imagery of a river flood in northwestern Iran to classify flooding. A portion of the imagery was manually labeled into five landcover classes (including water) and used to train a multilayer perceptron neural network. The study achieved a classification accuracy of 70% using manually labaled reference images for validation. Classification accuracies can often be improved by including features, such as textural metrics or spectral band indices, in addition to the original spectral bands of the imagery. Feng *et al.* (2015) used 0.2m resolution color imagery from a UAV to map urban flooding in Yuyao, China. The authors performed supervised classification with a random forest classifier using six gray-level co-occurrence matrix (GLCM) texture features, as well as the three spectral bands. Using the set of nine features, the authors obtained an overall classification accuracy of 87%. Malinowski *et al.* (2015) used multispectral, 2m resolution WorldView-2 imagery to perform supervised classification of floodwater in a river valley in Denmark. Rather than using a pixel-based approach, classification was performed on image objects, which were formed using a multilevel segmentation in the popular image classification software eCognition. Images were preprocessed to remove areas of high vegetation and shadows before classification. In addition to the spectral bands, several spectral indices, elevation data from lidar, and principal component bands were included in the classification. Using a decision tree classifier, the authors obtained an overall classification accuracy of 95%. Although supervised methods are able to achieve accurate classifications, especially when augmenting training data with additional features, they require a user to spend time to manually label data. Manually labeling training data can be a slow and labor-intensive process, which is not ideal in an emergency response situation where management officials need maps in near realtime to make administrative decisions.

Some unsupervised methods use change detection between a pre-flood and post-flood image to identify flooding. Wang *et al.* (2002) used a pair of Landsat 7 multispectral, 30m images from before and after a flood event in North Carolina to map flooding. Pixels were classified as water or non-water in each image using the value of the sum of a NIR and midwave infrared (MWIR) spectral band. Pixels that were classified as non-water in the pre-flood image and water in the post-flood image were then considered flood pixels. No quantitative results were published, but the authors state that most flood regions were confirmed

to be correct through visual inspection. Similarly, Ogashawara *et al.* (2013) used a pair of Landsat 5 multispectral, 30m images to classify flood regions near the upper coast of Texas, caused by Hurricane Ike. The Normalized Difference Water Index (NDWI) was calculated for the pre- and post-flood images using the NIR and SWIR bands. Flood regions were identified by thresholding the difference between the two NDWI images. An accuracy of 86% was obtained using this method. However, the main drawback of change detection approaches is the need for an existing pre-flood image and the registration of the two images.

Approaches using only post-event images are typically water mapping techniques, rather than strict flood mapping techniques. For example, Xie *et al.* (2016) used a single 30m multispectral Landsat 8 image to perform surface water mapping in urban regions of China. The algorithm used the NDWI to identify pure land, mixed land/water, and pure water pixels. Spectral mixture analysis was performed on the mixed pixels and pixels with a high water fraction were used along with the pure water pixels to generate water maps. The study reported producer and user accuracies of 83% and 92%, respectively, for water pixels. Sun *et al.* (2012) used a target detection approach to classify water after in Binghamton, NY after Tropical Storm Lee using a 0.15 m resolution color image. A single flood pixel was selected by a user and the spectral vector of that pixel was used with spectral angle mapper (SAM) to identify the rest of the water pixels, based on spectral similarity. Although no quantitative results were presented, a visual inspection of the algorithm output suggests it is an effective method.

The review of the literature presented here shows that there are many approaches to flood mapping and that the best technique is dependent on both the input data and the application. The research presented in this chapter proposes a technique that implemented very high spatial resolution airborne imagery to automatically map flooding in roadways. In order to achieve full automation, the technique by extension must be unsupervised. Change detection is useful when trying to separate water pixels that are normally water (*i.e.*, river or stream pixels) from pixels that are normally land. For this research, we are interested in mapping flooded roadways, and because roadways are never supposed to be flooded, the requirement of a pre-flood image does not apply. The proposed method uses a signature-matched detection approach to find pixels spectrally similar to an automatically selected seed flood pixel. Without needing specific spectral

bands to compute spectral indices, the algorithm is more robust to different types of input imagery.

4.2.2 Depth Estimation

Once a flood map has been created, water depth can be calculated by subtracting the elevation of the land from the elevation of the water surface (Sanyal and Lu, 2004). The land elevation typically is assessed based on a high-resolution DEM, which is often created from airborne lidar. The source of the water elevation depends on the nature of the study. For example, Jones *et al.* (2001) used hydraulic models with historical data to simulate flooding conditions and estimate water level and flood extent. High spatial resolution elevation data were then subtracted from the estimated water level to estimate the flood depth. Ali *et al.* (2001) created flood maps using SAR imagery and then calculated water depth for individual flood cells using elevations from a DEM. The authors assumed that the water level was flat in localized regions and used the height of the DEM at the edge of the region as the water surface elevation. The water depths were calculated by subtracting the DEM elevations from the water surface elevation. The authors found their method to be accurate within 30cm by validating with stream gauge stations.

Our method generates a water map using target detection and then calculates the water depth by subtracting the DEM elevations of water pixels from the elevation of the land-water border. This approach takes advantage of the flatness of water in connected flood regions. The main contribution of the research in this chapter is a fully automatic algorithm for detecting flood pixels in a flood image (and equivalently, the road network) and estimating the depth of the water at those pixels using a DEM raster. The proposed method is validated using imagery from the severe Mississippi River flooding in January 2016. The rest of this chapter is organized as follows: Section 4.3 introduces the study area and data used for algorithm development and testing, Section 4.4 describes the proposed algorithm in detail, Section 4.5 discusses experimental results, and finally, Section 4.6 includes conclusions drawn from the research.

4.3 Study Area and Data

Heavy rainfall, attributed to a strong El Niño season, resulted in severe flooding of the Mississippi and Arkansas Rivers (United States of America (USA)) in early January of 2016. The improbable winter flood set record high water levels at several streamgages, in some instances by more than two feet (NASA, 2016). Communities along the river were impacted from Illinois, all the way south to Louisiana, USA. Dozens of casualties, over $1B in damages, and significant flooding of major highways (some stretches as long as 24 miles), and other roads made this one of the worst flooding events in the region in recorded history (NOAA, 2016).

The National Geodetic Survey (NGS) collected airborne imagery of the flood extent along the Mississippi River and Arkansas River between January 2-18, 2016. red, green, and blue (RGB) imagery from an Applanix DSS580 and NIR imagery from a PhaseOne were acquired simultaneously aboard a King Air aircraft, operating at an altitude of about 2,300m. The imagery was registered in post-processing and has a GSD of 0.5m (NOS, 2016). Over 4,900km^2 of flood imagery was collected, spanning all the way from St. Louis (Missouri, USA) to Baton Rouge (Louisiana, USA). For development and testing of the proposed algorithm, the portion of the imagery that fell within Missouri, USA, was used.

In addition to the imagery, DEM images and road polygons were used as inputs. The DEM images were sourced from the 3-D Elevation Program (3DEP) geospatial data repository. The DEMs were rasterized from airborne lidar data and have a GSD of approximately 1m. The road polygons were created from buffers around road centerlines in Maptitude® mapping software. Figure 4.1 shows a map of the study area, with footprints of the three data sources shown in different colors. Ten of the flood images, as well as the corresponding DEM images and road polygons, were used for testing and validation.

4.4 Methods

The proposed algorithm ingests an airborne flood image, a DEM image, and a set of road network polygons as inputs, detects floodwater in the image, and then determines the depth of the water in the roadways.

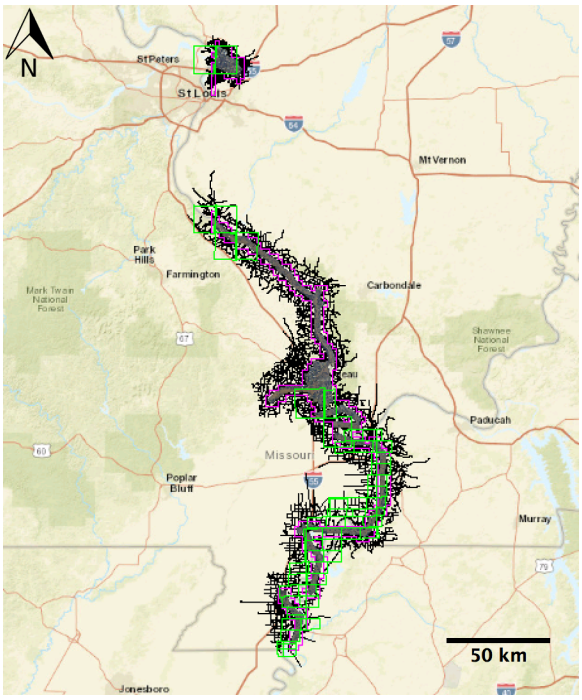


Figure 4.1: A map showing the data footprint in Missouri, USA. The flood imagery is outlined in pink, the DEM borders are marked in green, and the road network is marked in black.

The method can be divided into three stages: seed pixel detection, detection of roadway flood pixels, and estimation of water depth. A flowchart of the algorithm is presented in Figure 4.2.

4.4.1 Seed Pixel Detection

The first step of the algorithm is to find a single flood pixel, called a seed pixel, whose spectral information will later be used to detect the rest of the flood pixels in the input image, I. Floodwater appears mostly homogeneous in overhead imagery, so potential seed pixels are identified as those in homogeneous regions. Entropy can be used as a statistical measure of the amount of information present in a given region (Haralick, 1979). Entropy is computed for each pixel in the NIR band of the image using a 9x9 pixel moving window. The NIR band is used because NIR energy is strongly absorbed by water, making it appear dark and flat in the image (Klemas, 2014). If no NIR band is available (*i.e.*, if the image is only RGB), then the RGB bands are converted to the hue, saturation, and value (HSV) color space. The HSV color space is a cylindrical-coordinate representation of points in which the angle around the central axis is the hue, the distance from the center is the saturation, and the distance along the axis is the value (Sural *et al.*, 2002). Floodwater appears particularly uniform in the hue band due to its homogeneous color, so the hue band is used in place of a NIR band. The entropy image is binarized using Otsu's Method (Otsu, 1975) and inverted such that low entropy pixels (*i.e.*, flood pixel candidates) have a value of one and high entropy pixels have a value of zero.

In order to obtain as pure a flood pixel as possible, the candidate mask is cleaned of small, disconnected regions. The small regions are typically caused by the presence of trees, waves, or other small objects in the water. A morphological opening, followed by an area opening, is used to remove these regions from the mask. Figure 4.3b shows an example of a candidate mask after morphological processing.

At this point it is assumed that all of the remaining nonzero pixels in the candidate mask are flood pixels. The seed pixel is taken as the pixel with the greatest inverse black-white distance, or in other words, the nonzero pixel that is spatially farthest from any zero pixels. The assumption is that flood pixels farther from the edge will have less influence from nonflood pixels. Figure 4.3b shows an input image with the

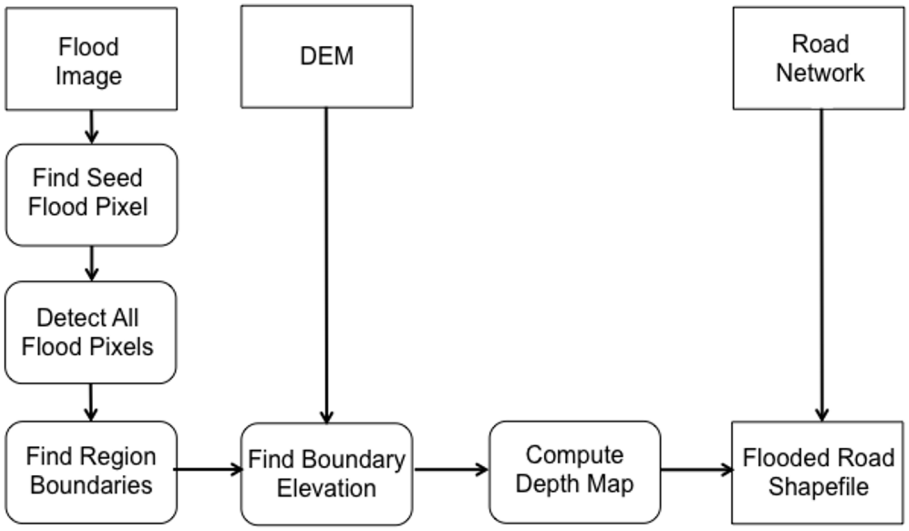


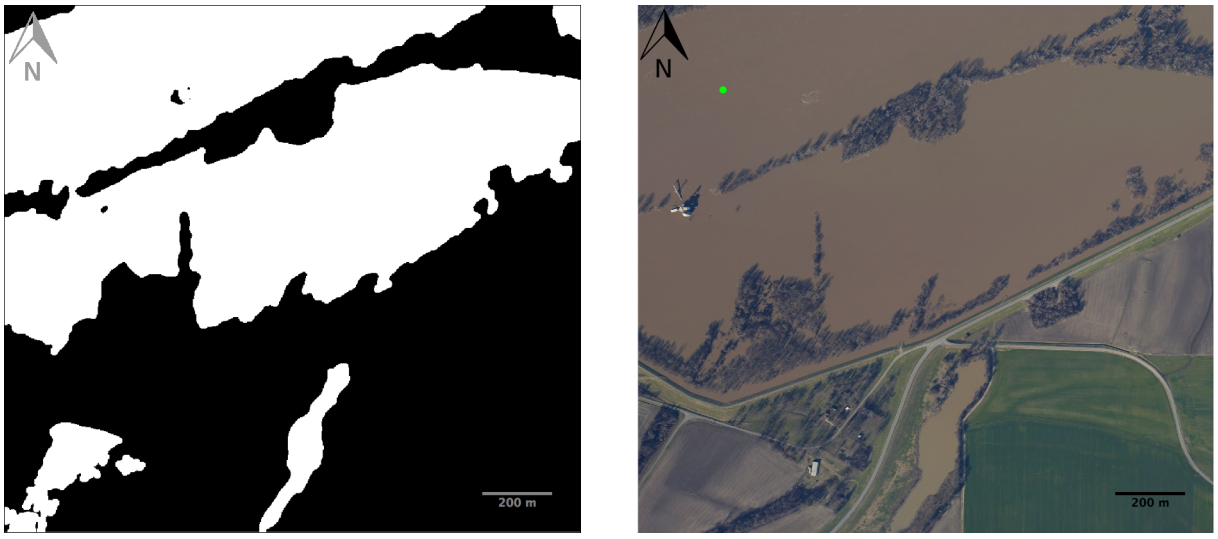
Figure 4.2: A flowchart of the roadway flood detection and water depth estimation workflow.

automatically selected seed pixel highlighted in green. The rest of the flood pixels are then found via target detection using the spectral vector of the seed pixel as a reference.

4.4.2 Detection of Flood Pixels

Roadway flood pixels are first detected by finding all of the flood pixels in the image, and then masking those pixels to the roadways using a road network shapefile. A target detection approach is used to find the flood pixels. The spectral vector of the seed flood pixel, $\mathbf{s}$, is used as the reference spectrum from which the other pixels will be detected. The dimensionality of the spectral vector is equal to the number of bands in the input flood image.

Many target detection algorithms make use of statistics of the reference, background, and noise (if known) pixel distributions. Particularly of interest are the mean, μ, and covariance matrix, Σ. In order to calculate meaningful values for these metrics for the flood spectrum, a distribution of pixels is required, rather than just a single seed pixel. Spectral Angle Mapper (SAM) is used to find pixels that are similar to the seed flood pixel. SAM is a signature-matched detection algorithm that uses the angle between the reference spectrum,



(a) (b)

Figure 4.3: A seed pixel is automatically selected from a set of candidate flood pixels: (a) The candidate flood pixels produced through entropy filtering and morphological processing; and (b) The input image with the seed pixel, shown as a green circle.

s, and the spectrum of another pixel, **x**, as a measure of similarity between the two pixels (Equation 4.1) (Eismann, 2012).

$$r_{SAM}(\mathbf{x}) = -\cos^{-1}\left(\frac{\mathbf{s}^T\mathbf{x}}{\sqrt{(\mathbf{s}^T\mathbf{s})(\mathbf{x}^T\mathbf{x})}}\right) \tag{4.1}$$

where r_{SAM} is the detection statistic (*i.e.*, measure of similarity) and smaller values imply that the two pixels are more similar.

The initial set of flood pixels is the set of pixels in the r_{SAM} image with a value below a threshold, T_{SAM}. A value of 0.0675 was chosen for T_{SAM} through analysis of receiver operating characteristic (ROC) curves created with manually labeled flood images. The mean ($\boldsymbol{\mu_s}$) and covariance matrix ($\mathbf{\Sigma_s}$) of the initial set of flood pixels is then calculated. Figure 4.4b shows the initial set of flood pixels detected using SAM for the input image shown in Figure 4.4a. Similarly, the mean ($\boldsymbol{\mu_b}$) and covariance matrix ($\mathbf{\Sigma_b}$) of the background pixels (*i.e.*, all pixels in the r_{SAM} image above T_{SAM}) are computed.

Now that information about the target and background spectra distributions are known, a more robust signature-matched detection algorithm can be used to detect the flood pixels. The quadratic spectral filter (QSF) is an algorithm that is robust to false alarms, because it only becomes large when the evaluated spectrum is both statistically distant from the background and similar to the reference (Eismann, 2012). The QSF is calculated using Equation 4.2:

$$r_{QSF}(\mathbf{x}) = (\mathbf{x} - \boldsymbol{\mu_b})^T {\boldsymbol{\Sigma_b}}^{-1} (\mathbf{x} - \boldsymbol{\mu_b}) - (\mathbf{x} - \boldsymbol{\mu_s})^T {\boldsymbol{\Sigma_s}}^{-1} (\mathbf{x} - \boldsymbol{\mu_s}) + \log \frac{|\boldsymbol{\Sigma_b}|}{|\boldsymbol{\Sigma_s}|} \tag{4.2}$$

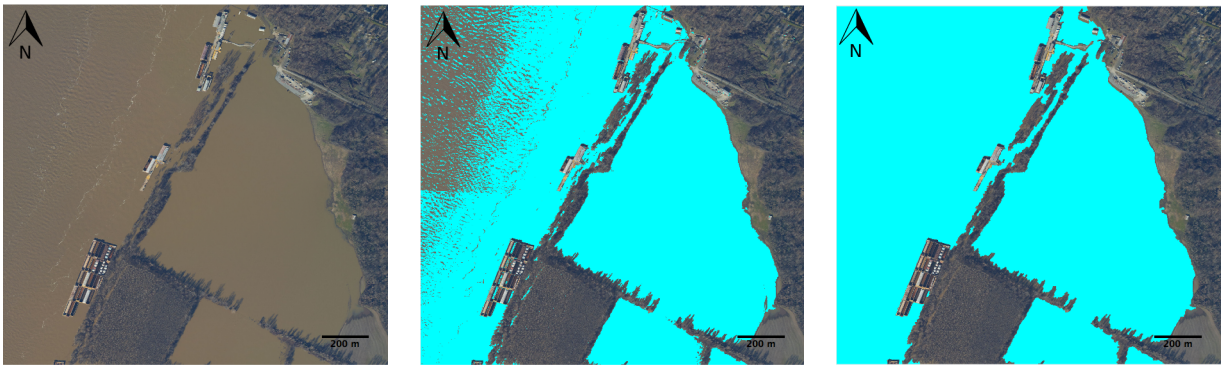
where r_{QSF} is the detection statistic. The original equation includes terms for noise, but noise is ignored in the proposed method, because insufficient data are available to characterize it.

As with SAM, the set of flood pixels is determined using a threshold, T_{QSF}. To improve the spatial integrity of the flood map and remove small sections of pixels, a morphological opening, followed by an area opening, is again applied. Figure 4.4c shows the final set of flood pixels detected using QSF for the input image shown in Figure 4.4a.

4.4.3 Estimation of Water Depth in the Roadways

One of the main objectives of this research is to determine the depth of water in the roadways. Towards that goal, the flood map generated in the previous step is divided into connected components in order to find regions that are in contact with roads. Figure 4.5 shows a set of road polygons, overlaid on top of a flood map, with connected components presented in different colors. In that image, of the eleven individual flood regions, only one component (the biggest one, shown in blue) actually abuts the road and will be used for water depth estimation. Flood regions that don't touch roads are discarded and the remaining regions are processed one at a time.

The main assumption for water depth estimation is that because water is level, the elevation where the water meets the land is the elevation of the water for the entire flood region. A DEM is aligned to the flood image and linearly interpolated to match its resolution, resulting in an elevation value for every pixel of the



(a) (b) (c)

Figure 4.4: Flood pixels are obtained using two stages of signature-matched detection: (a) The original flood image; (b) The initial set of flood pixels detected using SAM; and (c) The final set of flood pixels detected using QSF.

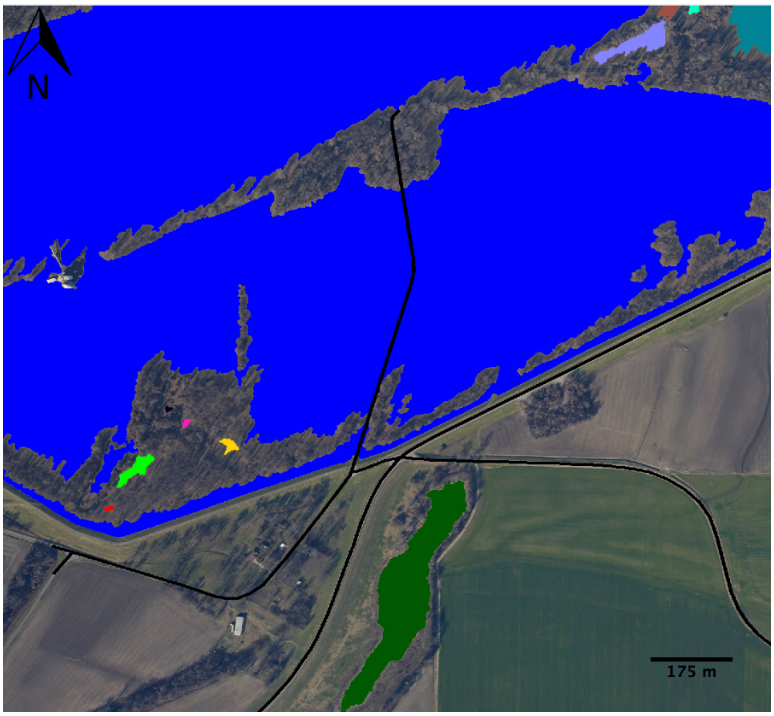


Figure 4.5: Connected components (each component is shown in a different color) and the road network are overlaid on a flood image. Only components that the road network passes through are used for water depth estimation.

flood image.

To estimate the water depth for the flood pixels in a flood region, the indices of the boundary pixels are used to retrieve the land elevations at the edge of the water from the DEM. Theoretically, based on the flat water assumption, all of the elevations should be the same. However, in practice, the actual elevations vary due to several factors. Water boundaries due to objects such as boats, bridges, and trees lead to boundary elevations that are much lower than the actual land-water boundary. Therefore, the maximum elevation of all the boundary elevation values is used as the elevation of the floodwater surface. Figure 4.6 displays a flood image with the boundary pixels plotted in black and the boundary pixel with the highest elevation highlighted in green.

After the water surface elevation, e_s, has been obtained, the height of the water, h_w, for the rest of the pixels is calculated by simply subtracting the value of the DEM at every pixel (equation 4.3).

$$h_w(x, y) = e_s - DEM(x, y) \tag{4.3}$$

where x and y are the matching pixel coordinates of both the water depth map and the DEM.

Shadows, vegetation, and other obscurations lead to omitted roadway flood pixels. Road pixels with elevations below the corresponding estimated water elevation, e_s, that are not classified as flood pixels, are labeled as candidates to be filled. A hole filling operation that fills omitted flood regions in the roadway that are fully bounded by flood pixels and have an area $\leq$ 2,500m^2 (which is approximately equivalent to 10,000 pixels in our imagery) is applied. The inundation map and depth image then are both updated to include the filled pixels.

Figure 4.7 shows a before and after image of a flooded region in Missouri, USA, along with a water depth map. The water depth map can easily be masked to the roadways using the road polygons.

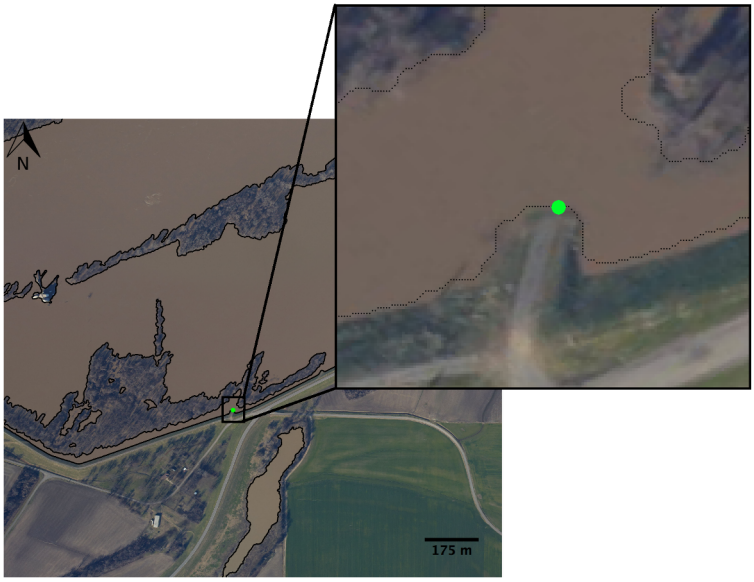
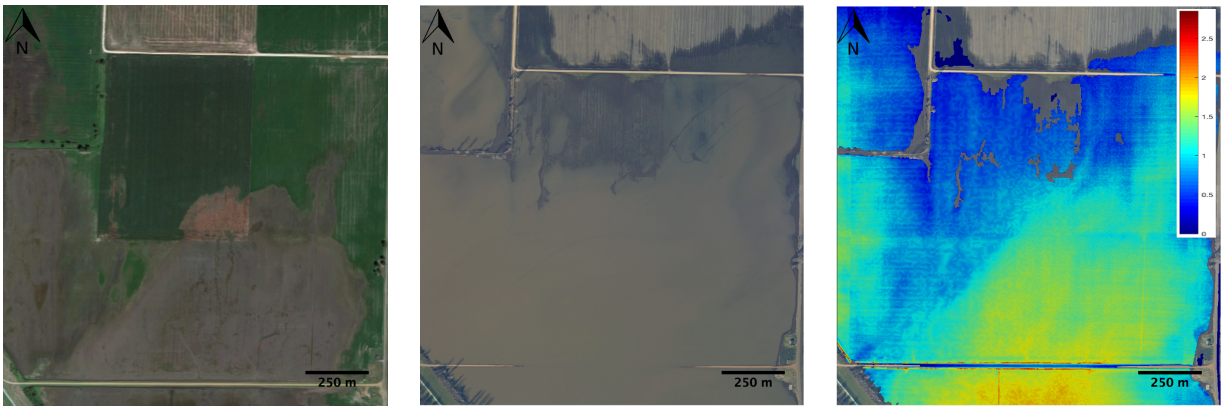


Figure 4.6: All of the flood region boundary pixels are plotted over the flood image in black. The pixel at the highest elevation is highlighted in green. The value of the DEM at this pixel is chosen as the flood surface elevation.



(a) (b) (c)

Figure 4.7: A flooded region in Missouri: (a) Pre-flood image (from Google Maps); (b) Post-flood image; and (c) Water depth image (the units on the scale are meters).

4.4.4 Experimental Validation

4.4.4.1 Validation of Roadway Flood Pixel Detection

To our knowledge, no maps of roadway flooding were ever created for the fairly recent Mississippi River, USA, floods of 2016. Therefore, reference data were created by manually labeling flood pixels in the NGS imagery that fall within the road network polygons. This process was carried out for all ten of the flood images used for testing. Figure 4.8 shows the labeled flood image for scene four, with the manually traced flood pixels colored red.

Analysis of roadway flooding detection was performed by comparing the flood map output and the manually digitized flood polygons. A pixel-by-pixel comparison was carried out, and each flood pixel was classified as either a TP, FP, or FN. Three metrics were calculated to characterize flood mapping performance. *Completeness* is a measure of errors of omission and reflects the percentage of roadway flood pixels in the reference data that were detected by the algorithm (equation 4.4). *Correctness* is a measure of errors of commission and reflects the percentage of flood detections by the algorithm that are true flood pixels (equation 4.5). *Quality* measures the overall performance of the flood detection algorithm and takes into account both the completeness and the correctness of the results (equation 4.6) (Heipke *et al.*, 1997).

$$Completeness = \frac{TP}{TP + FN} \times 100, \tag{4.4}$$

$$Correctness = \frac{TP}{TP + FP} \times 100, \tag{4.5}$$

$$Quality = \frac{TP}{TP + FP + FN} \times 100 \tag{4.6}$$

The same roadway flood detection analysis was performed using only the RGB spectral bands of the imagery in order to investigate the ability of the algorithm to operate without a NIR band. The same ten images were used and the NIR band was simply discarded.

Figure 4.8: The flood image for scene four, with the manually digitized roadway flood polygons overlaid in red.

4.4.4.2 Validation of Water Depth Estimation

There was no direct manner in which to validate the water depth estimation without field surveys of water depth collected around the same time that the flood imagery was collected. However, if the pixel from which the flood boundary elevation was obtained was correct, then the uncertainty of the water depth estimation should theoretically be equivalent to the accuracy of the DEM. Therefore, to obtain a lower bound on the error in the inundation depth estimation, the automatically detected water-land border was compared to a manually selected one, and the difference in elevation was computed. Fifteen border locations were analyzed for both the flood maps generated from the RGB-NIR imagery and those generated from the RGB imagery.

4.5 Results and Discussion

The proposed workflow was applied to ten flood images, DEM rasters, and road networks in Missouri, USA. The extent of the flooding in the roadways was determined using signature-matched detection and

morphological processing, and the corresponding water depths then were calculated using the water surface elevation obtained from the DEM at the land-water border.

4.5.1 Roadway Flooding Detection Results

4.5.1.1 RGB-NIR Imagery

The results of the RGB-NIR flood mapping analysis are displayed in Table 4.1. Overall, the flood detection algorithm was successful in detecting roadway flood pixels, achieving a quality score of nearly 97% across the ten scenes. Completeness and correctness scores of approximately 98% and 99%, respectively, suggest that false negatives and false positives were minimal. The flood detection results for one of the test images is displayed in Fig. 4.9.

4.5.1.2 RGB Imagery

The results of the flood detection using only RGB bands are presented in Table 4.2. As expected, the algorithm obtained less accurate results without the NIR band, obtaining a quality score of 69.21%. The main reason for the low quality score was the large number of false positives, resulting in a correctness of 70.96%. However, the algorithm was still able to achieve a completeness of 96.56%, meaning that most of the true flood pixels were accurately classified. The flood detection results for one of the scenes that had many false positives (34,714 pixels) is shown in Figure 4.10.

4.5.2 Roadway Water Depth Estimation

4.5.2.1 RGB-NIR Imagery

Fifteen border locations were analyzed for the flood maps produced using RGB-NIR imagery. The automatically detected border was on average 2.01 ± 0.11m displaced horizontally from the manually selected location. The 2m average border displacement corresponded to an average depth error of 0.08 ± 0.08m based on the DEMs of the test scenes. The roadway inundation depth map for one of the test scenes is presented

Table 4.1: Roadway flooding detection results for all ten scenes. TP, FP, and FN are numbers of pixels while C_{FN}, C_{FP} and Q are percentages.

Image	TP	FP	FN	C_{FN}	C_{FP}	Q
1	117,052	577	1,698	98.47	99.51	98.09
2	131,047	490	1,874	98.59	99.63	98.23
3	34,559	671	0	100.00	98.10	98.10
4	37,858	1,154	1,263	96.77	97.04	94.00
5	22,595	0	0	100.00	100.00	100.00
6	64,351	338	691	98.94	99.48	98.43
7	68,283	0	1,117	98.39	100.00	98.39
8	45,470	178	28	99.94	99.61	99.55
9	79,519	761	4,818	94.29	99.05	93.44
10	111,377	1,936	5,190	95.55	98.29	93.99
Total	712,111	6,105	16,679	97.71	99.15	96.90

Figure 4.9: Flood mapping results for a section of image six. While a majority of roadway flood pixels were correctly classified, false positives exist near the silo to the left and some false negatives are present at the flood-road boundary.

Table 4.2: Roadway flooding detection results for all ten scenes using only the RGB spectral bands. TP, FP, and FN are numbers of pixels while C_{FN}, C_{FP} and Q are percentages.

Image	TP	FP	FN	C_{FN}	C_{FP}	Q
1	110,451	18,783	8,299	93.01	85.47	80.31
2	130,851	11,723	2,070	98.44	91.78	90.46
3	34,529	95,411	30	99.91	26.57	26.57
4	38,558	1,914	563	98.56	95.27	93.96
5	22,595	8	0	100.00	99.96	99.96
6	64,756	34,714	286	99.56	65.10	64.91
7	67,991	0	1,409	97.97	100.00	97.97
8	45,498	61,515	0	100.00	42.52	42.52
9	82,832	62,369	1,505	98.22	57.05	56.46
10	105,677	1,522	10,890	90.66	98.55	89.47
Total	703,738	287,989	25,052	96.56	70.96	69.21

Figure 4.10: Flood mapping results for a section of image six using only the RGB spectral bands. While almost all true flood pixels were correctly classified, many false positives were detected near the silos and where the water meets the road.

in Fig. 4.11.

4.5.2.2 RGB Imagery

The same fifteen border locations were analyzed for the RGB flood maps and the results showed that the automatically detected border was on average 2.89 ± 0.14m displaced horizontally from the manually selected location. The 2.9m average border displacement corresponded to an average depth error of $0.14 \pm .13$m based on the DEMs of the test scenes.

4.5.3 Discussion

Analysis of the results revealed two main causes of misclassification of flood pixels. False positives were caused by wet portions of road that had been previously inundated before the water began to recede. False negatives were dominated by small inundated regions of roadway near the flood boundary that had an elevation slightly above the estimated water level. The underestimation of the water level in some cases was caused by uncertainty in the detected land-water pixel, which was caused by omission of flood pixels at the land-water border, due to their spectral differences compared to pure flood pixels. Future work needs to

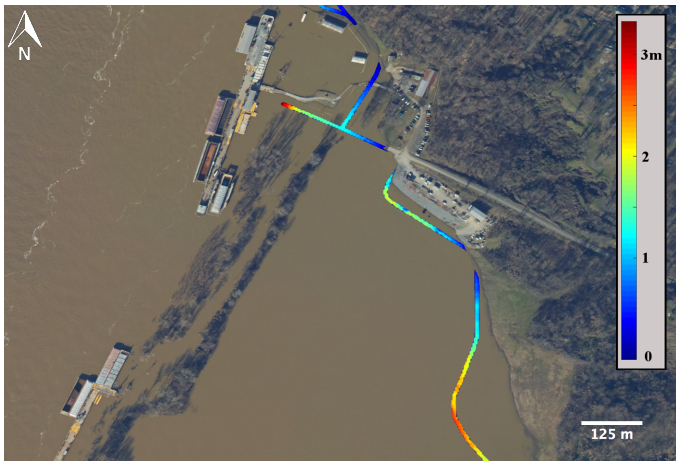


Figure 4.11: The estimated inundation depths for flooded roadways of scene four.

address these issues by treating pixels near the land-water border as mixed pixels and performing further spectral analysis on them to obtain a final classification.

The use of the 3DEP DEMs and Maptitude® road centerline data introduced some unavoidable sources of error into the algorithm. For example, the DEM resolution was twice as coarse as that of the imagery, and consequently the estimated water depths were degraded. Additionally, a fixed-width buffer around the road centerlines was used to create the road polygons, which did not accurately represent the varying widths of the different roads in the test scenes. These data sources were the best available during this research, but the availability of finer DEMs and true roadbed polygons in future work would help yield more accurate results.

It was clear from the results that the algorithm achieved much better performance on RGB-NIR imagery as compared to the same imagery using only the RGB bands (96.9% vs. 69.21%). This was not surprising, since water was the most easily identifiable in the NIR band where it appeared very dark due to the absorption of NIR energy. However, it is worth noting that although the overall performance was poor on the RGB imagery, the completeness scores were comparable (97.71% for RGB-NIR and 96.56% for RGB). These results suggest that even when the NIR band was not available, the algorithm was still able to detect most of the flood pixels, but it also misclassified many non-flood pixels because they were spectrally similar in the

visible portion of the spectrum. This algorithm was not intended to be used for 3-band imagery, but our experiments show that it can still be used at the expense of introducing more false positives.

It is important to compare the performance of our algorithm with other high performing flood inundation estimation methods in order to understand its strengths and weaknesses. One of the common approaches in the literature was to train a supervised classifier using spectral bands, as well as other features derived from the imagery. In Feng *et al.* (2015), RGB spectral values, as well as Haralick texture features from 0.2m UAV imagery, were used to train a random forest classifier resulting in 87% accuracy. Similarly, Malinowski *et al.* (2015) used eight spectral bands along with spectral indices to train a decision tree classifier, which obtained a classification accuracy of over 95%. Although the Feng *et al.* (2015) approach performed well on 3-band imagery, the added burden of manually generating training data in these two algorithms, as well as their slightly lower accuracies, leads us to believe that our method is a better choice for emergency response situations. Ogashawara *et al.* (2013) used change detection with pre- and post-flood NDWI images derived from Landsat-5 and obtained an accuracy of 86%. Comparatively, our method does not require pre-event images, and was able to obtain an accuracy of over 96%.

Finally, the flood mapping method using SAM in Sun *et al.* (2012) was a strong approach, in fact, it was part of the motivation for our research. We believe we have improved upon that method by introducing automatic seed pixel detection, thereby eliminating the need for user interaction. Additionally, the introduction of the QSF helped to reduce false alarms that occur when using only SAM. Although Sun *et al.* (2012) reported good performance through visual inspection, they did not provide any quantitative results for comparison.

4.6 Conclusions

In this chapter we proposed a fully automatic workflow for mapping flooded roadways and estimating water depth using four-band aerial imagery, DEM rasters, and road network polygons. The entropy of the NIR band was used to detect a seed pixel, while the SAM and QSF detectors were used to infer flood pixels in

the image. Water depth at flood pixels was determined by detecting the highest pixel at the land-water border and subtracting the DEM elevations of all the flood pixels. The flood mapping method achieved an overall quality of 96.9% when evaluated with high resolution imagery collected by the NGS during the severe flooding event of the Mississippi River, USA, in January 2016. A limiting factor in the performance of the algorithm was the ability to locate the pixel at the exact land-water border, which resulted in some omitted flood pixels and an uncertainty in the inundation depth estimation of at least 0.08m.

Given the results obtained by this research, we believe our method could be of great value during flooding to automatically process airborne flood imagery and provide maps of roadway inundation and water depth. In using an unsupervised, monotemporal approach, we eliminate the need for significant human interaction and pre-event imagery, thereby reducing potential delays in the emergency response. Furthermore, our method could be used in a predictive manner to estimate water depths for roads or other regions. Given a simulated water surface elevation, a flood mask could be generated from the DEM raster. The flood mask would then be processed as described in Section 4.4, resulting in a predicted water height for every pixel.

Chapter 5

Building Damage Assessment using Airborne Lidar

5.1 Introduction

Building damage assessment represents an urgent response priority following a natural disaster. This is true because determination of the damage status of buildings allows first responders to be directed to the most important locations, and resources, which are often a limiting factor in emergency response, can be utilized to their full potential. The geographic extent of large-scale disasters and potentially unsafe ground conditions often prohibit field-based assessments from providing a rapid overview of damaged regions (Eguchi *et al.*, 2008). Instead, remote sensing platforms can be utilized to collect 2-D or 3-D data over the affected area, which can in turn be used to assess damage, either manually by analysts or automatically by algorithms.

Much research has been conducted in building damage assessment using a wide variety of remotely sensed data. In the 2-D domain, optical and SAR imagery have been used to detect and classify the damage of buildings after disasters such as earthquakes. For optical imagery, features such as spectra, texture, shape, and building shadows have been used to detect damage in both change detection and post-event only analysis

(Dong and Shan, 2013). Change detection of optical imagery requires precise registration between the two sets of imagery and can result in false alarms due to illumination and color differences. Methods using SAR imagery typically exploit backscattering intensity and phase information to locate damage (Voigt *et al.*, 2011). Success with SAR data in urban areas has been limited due to issues arising from an oblique viewing geometry, occlusions, and multiple scattering from tall buildings (Dong and Shan, 2013). Some building types are also undetectable in the 2-D domain due to a lack of height information, such as "pancake collapses", in which one or more stories collapse onto themselves because of structural failures (Rehor *et al.*, 2008).

3-D data, typically in the form of point clouds, provide accurate height information that facilitates the detection of building damage. Geometric features like planarity and inclination angle, combined with surface features such as curvature and size, provide insight to the structural condition of a building. Point clouds can be collected using airborne lidar or formed during post-processing from imagery. Airborne lidar is advantageous for rapid building damage detection for several reasons. The data can be collected day or night and does not suffer from illumination shadows cast by tall buildings. Additionally, the data can be used as soon as it is downloaded from the sensor without the need for extensive post-processing (as compared to image-based point clouds), which is critical in a time-sensitive scenario such as disaster response (Baltsavias, 1999). Some research has used co-registered lidar data and optical imagery for building damage assessment using object-based image analysis, and obtained improved results over using just lidar data (Haiyang *et al.*, 2010; Ural *et al.*, 2011). Although these methods are effective, they need additional processing to co-register the sensing modalities and rely on a collection platform that can collect both modalities simultaneously or require multiple data collections. Other studies have performed change detection on pre- and post-event planar segments, derived from 3-D data, to classify damage based on changes such as volume reduction and inclination change (Schweier *et al.*, 2004; Rehor, 2007). The challenge with change detection is that it relies on the availability of pre-event data, which is not always the case. Even in cases where the pre-event data are available, precise registration is needed to avoid introducing damage false alarms, which can be difficult with two 3-D datasets that are often collected by different sensors, and at different point densities. Methods

that use only post-event lidar point clouds stand alone and can be applied to the raw point cloud data as soon as it is available from the laser scanner. These methods fit the time-sensitive criteria of a natural disaster response plan and will be the focus of this chapter. In the next section, a review of the methods in the literature that utilize only post-event airborne lidar point clouds for building damage assessment is presented.

5.2 Literature Review

A wide variety of techniques for detecting building damage from airborne lidar point clouds are present in the literature. All of the methods are based on extracting features and classifying damage at either the point, segment, or roof level. One of the main distinguishing features is the use of supervised classification versus unsupervised classification.

Methods that use supervised classification require a human operator to manually select labeled data to train a classifier. For the case of building damage assessment, this means supplying the classifier with training data with damage labels. For example, Rastiveis *et al.* (2015a) used a linear Support Vector Machine (SVM) to classify a DSM, rasterized from a point cloud collected after an earthquake. The SVM was used for binary classification of pixels into a debris class or intact class, based on eight texture features and height above ground. Buildings containing 30% or more debris pixels were classified as damaged. The study reported an overall classification accuracy of 91.6% on a dataset consisting of 43 buildings, of which only five were damaged. Elberink *et al.* (2011) performed supervised classification of planar segments, extracted from post-earthquake lidar data, using a rule-based classifier and a maximum entropy classifier. Five features were computed for each segment that represented its size, height, and planarity. The rule-based and maximum entropy classifiers obtained overall qualities of 56% and 60%, respectively. Similarly, Khoshelham *et al.* (2013) classified planar segments into damaged or undamaged using three different classifiers: a linear discriminant classifier, linear SVM, and random forests. An initial set of 18 segment features was reduced to a subset of six using forward selection and backwards elimination. Notable features from the subset included the ratio

of unsegmented points to segmented points, sphericity, and the height above ground. The authors were able to obtain an overall accuracy of 85% on a test set of 698 labeled segments, but did not supply a method for deriving building-level damage from the classified segments. The main limitation of damage assessment techniques based on supervised classification is that they require a user to manually label training data. As the number of features increases, the number of training samples needed to produce an accurate classification also increases. The training time can create a large bottleneck in the damage assessment workflow, which could negatively impact the response effort.

Unsupervised classification techniques found in the literature can be divided into two distinct groups, those that classify damage at the point level and those that classify damage at the building level. Labiak *et al.* (2011) proposed a method in which damaged roof points were detected by comparing the slopes of the lines formed with neighboring points. If the difference in slopes in both x- and y-directions was greater than some threshold, the point was considered damaged. Buildings containing 16% or more damaged points were classified as damaged. The method achieved an overall accuracy of 73.4%, but was only able to correctly classify 38.5% of the heavily damaged buildings in the dataset. In a continuation of that study, Labiak (2011) developed an improved damage assessment technique that used surface normals to identify damaged points. Building roofs were gridded into small tiles and the points within each tile were used to compute a surface normal. A histogram of the angles between the surface normals and a zenith vector was computed and the points in tiles with angles that fell in bins containing less than 20% of all the angles were considered damaged. A damage percentage was computed as the ratio of damaged points to total points in a building. Buildings with a damage percentage of 51% or higher were classified as damaged. The study achieved an accuracy of 68.3% on a dataset of 160 buildings, with many false alarms resulting from hipped roofs. Aixia *et al.* (2016) also used the angle between surface normals and zenith vectors to identify damaged points. A surface normal and angle was computed for every point using its nine nearest neighbors. By examining the angle distributions for buildings labeled as collapsed, partly collapsed, or undamaged, angular thresholds were derived to classify points in flat roofs and inclined roofs as damaged. The ratio of the angle standard deviation to the mean absolute deviation was used as an indicator of the severity of damage. Although

visual comparisons are presented for buildings with differing levels of damage, no classification results are presented in the paper.

Other damage assessment methodologies compute building-level features in order to make a more direct classification. For example, Yonglin *et al.* (2010) classified buildings as damaged if their inclination angle was above a certain threshold. The inclination angle was defined as the angle between the geometric axis of the building (the normal vector of the roof plane for flat buildings, or the sum of the normal vectors of "main" planes for buildings with slanted roofs) and the normal vector of the terrain the building sits on. However, classification results were only presented for a single undamaged and damaged building. He *et al.* (2016) proposed a damage classification method using a novel 3-D shape descriptor for buildings. Each building was represented as clusters of contours within a contour tree, and the shape descriptor was computed based on shape similarities within the contours. Buildings were classified as damaged if the shape descriptor met a set threshold. The algorithm was able to correctly classify 87% of the 1,875 buildings tested.

Despite the wide range of damage assessment mechanisms present in the literature, the majority of them share similar strategies. For example, planarity, surface normals, and the angles between surface normals and a vertical zenith vector appear in many of the previously mentioned publications. This is because building damage assessment lends itself naturally to an investigation of planarity and surface normals. A manmade surface is typically dominated by planes, but this assumption is often violated when buildings are damaged (Vosselman *et al.*, 2004). Roof points at high or near vertical angles are unusual for most intact buildings and can be used as an indication of damage. Combining aspects of work already in the literature, this chapter proposes an automatic building damage assessment methodology that identifies candidate damaged buildings as those with roof points at high angles and then performs a rule-based classification dependent on the planarity and height above ground of roof features. The objective of this research is to assess the ability to create an end-to-end, robust building damage assessment algorithm that is fully unsupervised and requires only a post-disaster point cloud. Based on the work presented in the literature, it is hypothesized that the research objective can be accomplished if appropriate point features are used as indicators of damage, and if building damage is adequately sampled in the lidar point clouds.

Section 5.3 introduces the dataset and building damage scale used for this research. Section 5.4 provides a detailed explanation of the methods used to preprocess the point cloud, detect buildings, and finally assess damage. In Section 5.5 experimental results are presented on airborne lidar data from the 2010 earthquake in Haiti, along with a discussion of the results, while conclusions and future work are addressed in Section 5.6.

5.3 Data

The study area used to test the proposed algorithm consists of seven sites located in the Haitian cities of Port-au-Prince and Carrefour, two regions that were heavily affected by the 7.0M_w earthquake that occurred on January 12, 2010 (see Figure 5.1). These sites were chosen because they contain both a wide range of building types and building damage types. Different construction types include: one- to three-story reinforced concrete buildings, masonry bearing walls, timber frames, and shanty housing made of reinforced concrete and masonry block with corrugated metal roofs (Rathje *et al.*, 2011). The damage level in buildings range from completely undamaged to fully destroyed, and everything in between.

The building damage assessment method proposed in this chapter requires only one input, an airborne lidar point cloud. The lidar data used for development and testing were collected on January 21, 2010 by Kucera International Inc. and RIT. The point clouds for the seven sites have an average point density of 4.2 pts/m^2 and were captured by a Leica ALS60 at an altitude of approximately 820m with a pulse rate of 150kHz. The vertical point measurement accuracy of the instrument is 0.15m. Multispectral imagery simultaneously was collected on the aircraft by the WASP system at a resolution of 0.15m (Messinger *et al.*, 2010). Although the imagery was not used for debris detection or volume estimation, it served as a reference for visualization and was used in figures throughout the paper. Figure 5.2 shows both a WASP image and lidar point cloud of one of the scenes from Port-au-Prince.

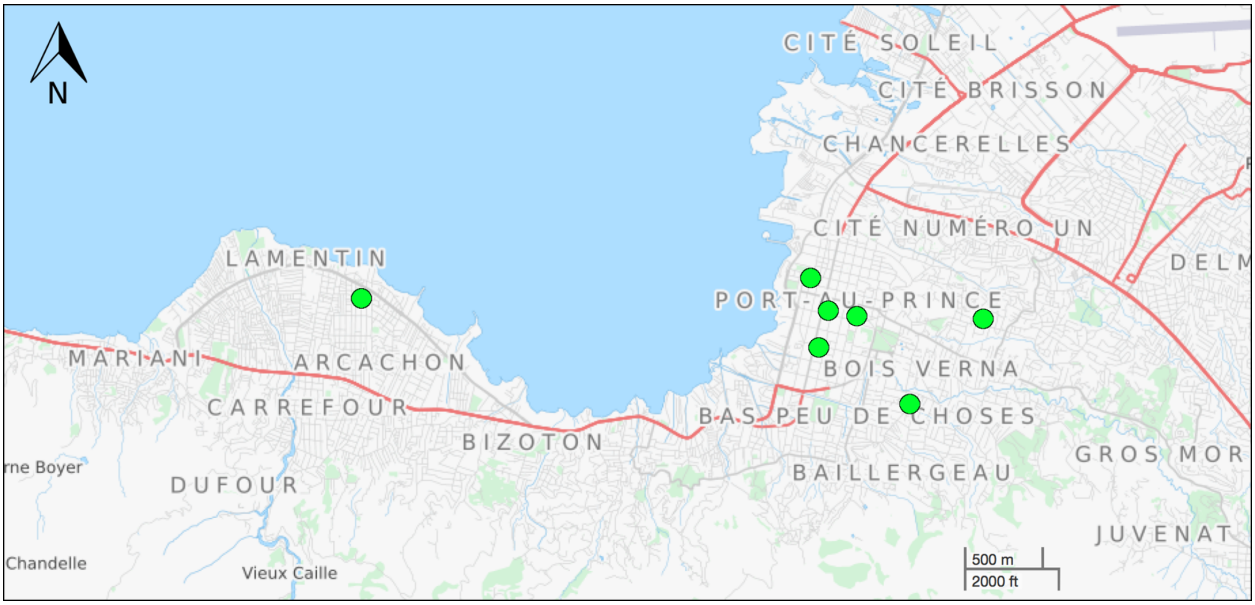
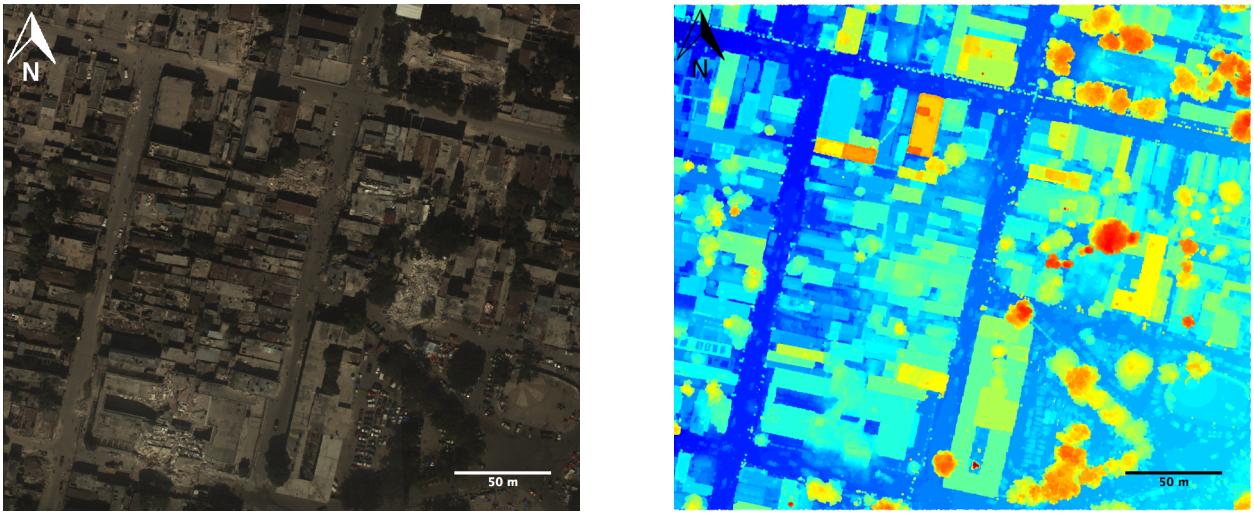


Figure 5.1: A map showing the portion of Haiti that was used in this research. The location of the seven test sites are marked by the circles.



(a) (b)

Figure 5.2: One of the sites from Port-au-Prince: (a) A WASP image of the scene; and (b) The corresponding lidar point cloud of the scene. The cloud is colored by height.

5.4 Methods

The proposed workflow takes an input airborne lidar point cloud, segments the cloud into individual building regions, and then classifies the damage level of the buildings. The method can be divided into three components: preprocessing, building segmentation, and damage classification. Figure 5.3 shows a visual representation of the workflow.

5.4.1 Point Cloud Preprocessing

The only input to the proposed workflow is an airborne lidar point cloud of a disaster-affected region. The required attributes for each point in the cloud are its x-, y-, and z-coordinates. The point cloud is filtered to remove noise and outlier points using a filtering technique called statistical outlier removal (Rusu, 2010). The statistical outlier removal algorithm first calculates the mean Euclidean distance, $\bar{d}$, between each point, p, and its k closest neighbors. Statistics of the mean distances are used to characterize the distribution across all of the points in the cloud. Specifically, the mean (μ_k) and standard deviation (σ_k) are calculated, and points that are considered statistical outliers, based on those two values, are removed (equation 5.1):

$$P^* = \{p \in P | \bar{d} \leq (\mu_k + \gamma \cdot \sigma_k)\} \tag{5.1}$$

where P^* is the entire point cloud after statistical outlier removal and γ is a scalar multiplier to control the severity of point removal.

Ground points are separated from nonground points (*i.e.*, buildings, vegetation, and vehicles) through an implementation of the PMF (Zhang *et al.*, 2003). The PMF works by using a morphological filter of gradually increasing window size, along with an elevation difference threshold, to separate ground and nonground points. The increasing window size is used to remove nonground objects of increasing size in each iteration (*i.e.*, small objects like bushes to start, and large objects like buildings in later iterations). One of the main parameters of the PMF is the initial elevation difference threshold, dh_0. Points that have a height above the estimated ground surface higher than dh_0 are classified as nonground points. By setting dh_0 to

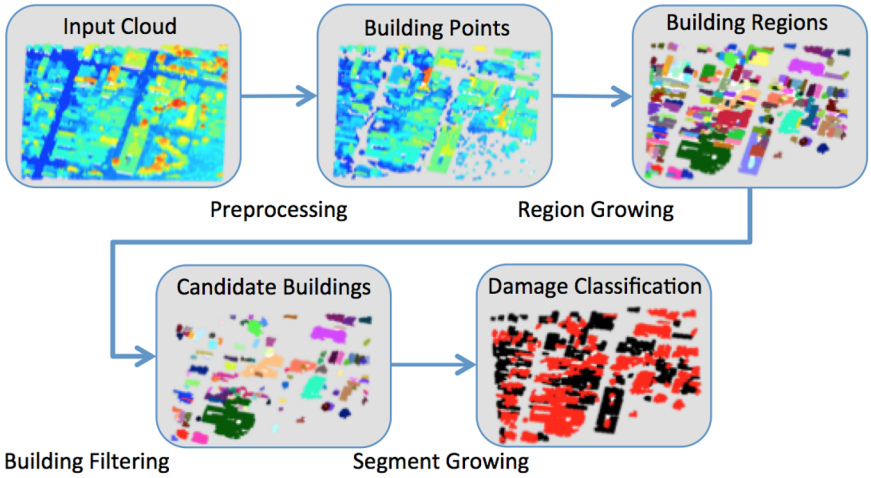


Figure 5.3: A flowchart of the building damage assessment algorithm.

be very low (we use the vertical accuracy of the lidar system, which is 15cm in our case), debris around the base of damaged buildings can be correctly classified as nonground, facilitating damage assessment in a later portion of the workflow. Following the PMF, a DEM is created by performing a Delaunay Triangulation with natural neighbor interpolation on the points classified as ground. A nDSM of the points classified as nonground is created by subtracting their elevations from the corresponding DEM elevations. The nDSM is stored for use later in the workflow.

The final step of preprocessing is vegetation removal. Vegetation contained in the point cloud can be mistaken for damaged buildings during damage assessment, so it is important to remove as much vegetation cover as possible during preprocessing. Vegetation removal is accomplished using a graph cuts optimization, based on local surface properties of points, originally proposed by Sun (2013). Manmade surfaces are typically locally smooth with little variation of surface normals in a small region. Vegetation, on the other hand, exhibits large variations of surface normals and high curvature. Therefore, local surface properties are used to distinguish between vegetation and nonvegetation points. First, the normal vector of each point, p, is calculated using all points in a local neighborhood, N_p, defined by the radius, r. The radius, r, is

automatically set as $2 \cdot \mu_k$. The neighborhood points of p are obtained using:

$$N_p = \{q \in P | d(p,q) < r\} \tag{5.2}$$

where q are all of the points in the entire point cloud P, and d is the Euclidean distance between two 3-D points. Eigenanalysis of the covariance matrix of N_p produces the eigenvalues $\lambda_1 < \lambda_2 < \lambda_3$. The eigenvector corresponding to λ_1 is the estimate of the point normal, $\vec{n}$. Airborne lidar data are collected from above, so the absolute value of the z-component of $\vec{n}$ is used to ensure the normal vector points outward from the surface. The curvature, v, of the point can also be computed using equation 5.3:

$$v = \frac{\lambda_1}{\lambda_1 + \lambda_2 + \lambda_3} \tag{5.3}$$

The neighborhood analysis is taken one step further by computing the distribution of normals in N_p. Eigenanalysis is again used, but this time on the covariance matrix of N_p, resulting in ${\lambda^n}_1 < {\lambda^n}_2 < {\lambda^n}_3$. The eigenvalue corresponding to ${\lambda^n}_2$ is representative of the variation of the local distribution of normals around the point. For simplified notation in equations, ${\lambda^n}_2$ will be represented as f.

A weighted graph, $\mathcal{G}$, of all of the points, $\mathcal{V}$, is constructed using each point and its four nearest neighbors. The weight on each link between two points is the inverse Euclidean distance between the two points. The energy function, E, which is comprised of a data term, D, and a smoothness term, S, is used for optimization, as shown in equation 5.4:

$$E(l) = \sum_{p \in \mathcal{V}} D_p(l_p) + \sum_{\{p,q\} \in N_p} S_{p,q}(l_p, l_q) \tag{5.4}$$

where l is the label assigned to a given point (*i.e.*, 1 for vegetation and 0 for nonvegetation).

The data term (equation 5.5) accounts for the curvatures and local normal variations for each point in the cost of making a cut:

$$D_p(l_p) = s_v \cdot e^{-\frac{v_p^2}{\sigma_v{}^2}} + s_f \cdot e^{-\frac{f_p^2}{\sigma_f{}^2}} \tag{5.5}$$

where s_v is a scalar coefficient for the curvature term, s_f is a scalar coefficient for the normal variation term, σ_v is the standard deviation of all of the point curvatures, and σ_f is the standard deviation of all of the normal variations.

The smoothness term (equation 5.6) controls the cost of neighboring points based on their labels:

$$S_{p,q}(l_p, l_q) = s_s \cdot \delta(l_p, l_q) \tag{5.6}$$

$$\delta(l_p, l_q) = \begin{cases} 0 & if \quad l_p = l_q \\ 1 & if \quad l_p \neq l_q \end{cases} \tag{5.7}$$

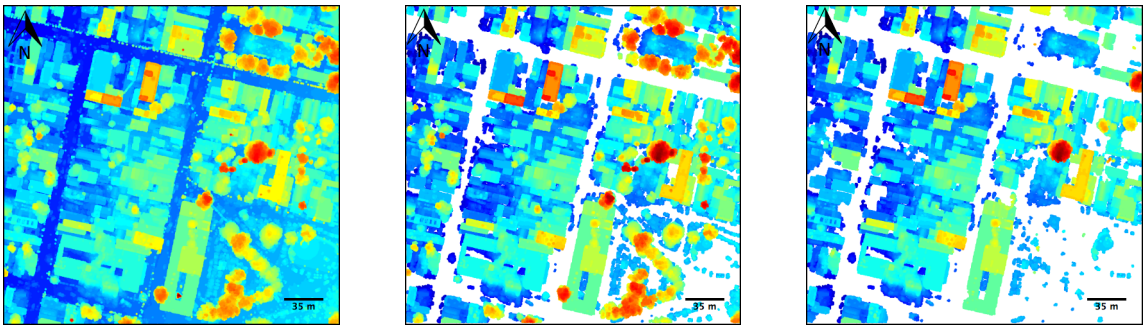
where s_s is a scalar coefficient.

As a result of the graph cuts optimization, all points are classified as either vegetation or nonvegetation. The points classified as vegetation are removed, which is the last step of preprocessing. Figure 5.4 displays an input point cloud at multiple stages of the preprocessing workflow.

5.4.2 Building Segmentation

After preprocessing, the remainder of the points consist of buildings, debris from damaged buildings, and small objects, such as vehicles. The next step is to segment both damaged and undamaged buildings into individual point clouds for damage assessment. To accomplish this task, a region growing with smoothness constraint approach is used (Rabbani *et al.*, 2006).

The inputs to region growing include all of the points (p), their normal vectors ($\vec{n}$), and their curvatures (v). The points are sorted in order of increasing curvature and the point with the lowest curvature is chosen as the first seed point. All of the points within a selected radius, r, of the seed point are considered as candidate points in the current region. If the angle, β, between the normal vectors of the seed point and

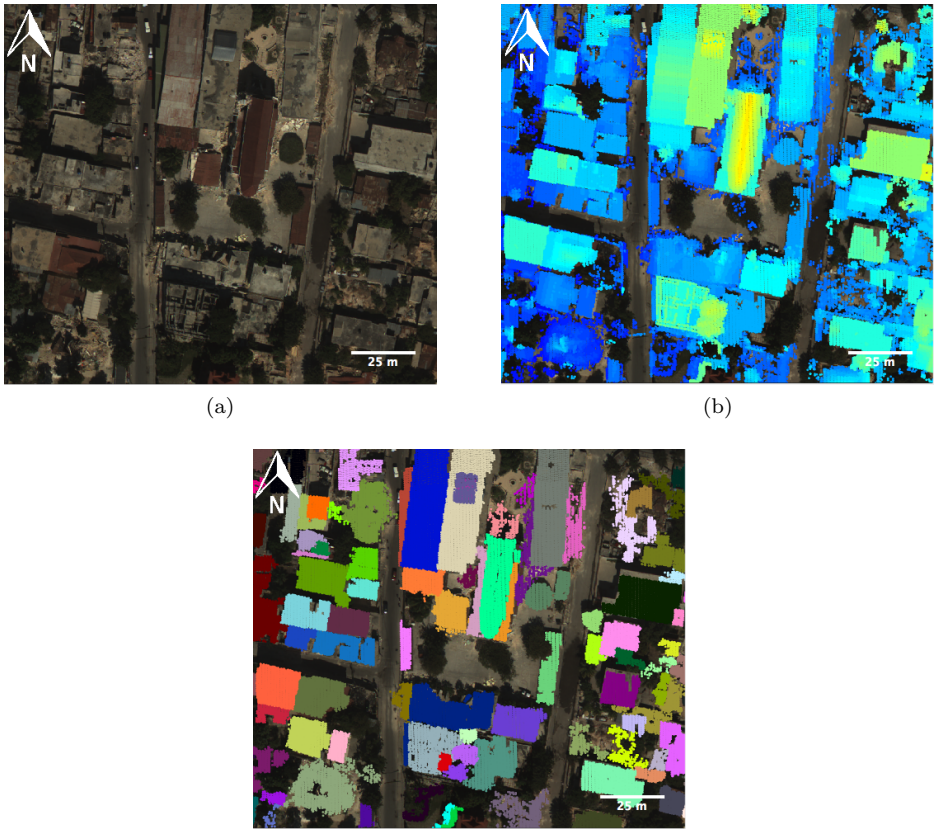


(a) (b) (c)

Figure 5.4: A point cloud from Haiti at various stages of preprocessing; the point cloud is colored by height: (a) The input cloud; (b) The point cloud after statistical outlier removal and removal of ground points; and (c) The point cloud after vegetation removal. Some trees remain in the point cloud (two can be seen in the top right), but a majority are removed.

candidate point are below an angular threshold, T_β, then the point is added to the current region. If the curvature, v, of the candidate point is below a threshold , T_v, it is added to a list of potential seed points. After all of the candidate points have been tested, the next seed point in the list is selected as the current seed point and the process is repeated. This sequence of steps is repeated until no more points can be added to the current region. The process is completed once all of the points have been assigned to a region. The value of r is automatically set to $2 \cdot \mu_k$.

The output of region growing is considered to be the set of buildings to be assessed for damage. A minimum point threshold, T_p, of 100 points is used to remove regions that are too small to be reliably considered for damage. Figure 5.5 shows an example of building segmentation of a point cloud that has already been preprocessed. Some undersegmentation is present as a result of liberal parameters ($T_\beta = 25^o$, $T_v = 0.05$) and the close proximity of buildings in the scene. These thresholds were determined through experimentation with several scenes. However, this undersegmentation allows debris to be segmented with buildings and helps to identify partially damaged buildings during the next step of the algorithm.



(a) (b)

(c)

Figure 5.5: A point cloud in a region of Port-au-Prince that was affected by an earthquake is segmented into buildings using region growing: (a) A WASP image of the scene; (b) The point cloud overlaid on the image. The point cloud is colored by height; and (c) The individual buildings resulting from region growing overlaid on the image. Each unique color represents one building.

5.4.3 Building Damage Assessment

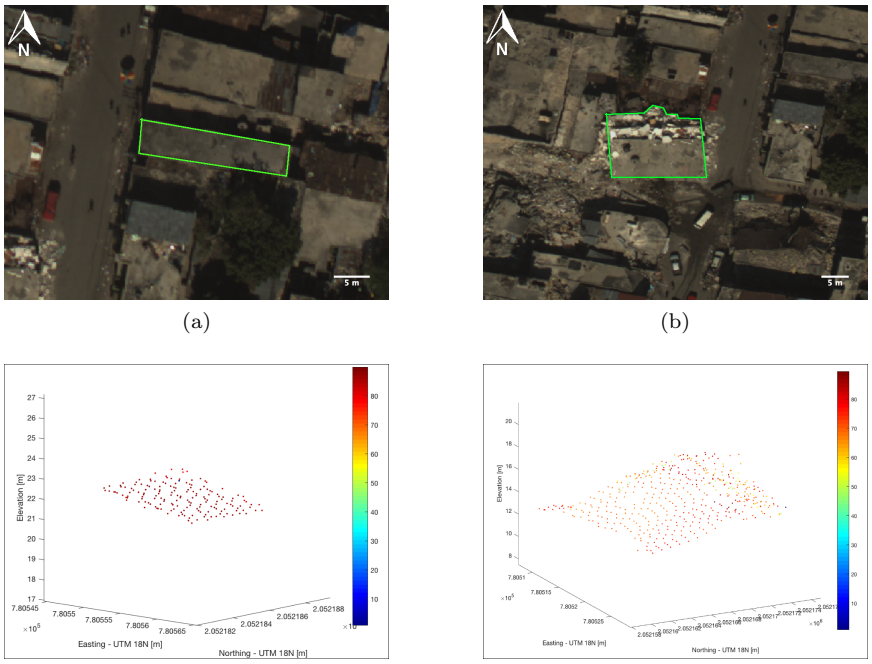
The proposed algorithm performs a classification of each building into one of two categories, damaged or undamaged. There is sometimes a large amount of segmented buildings in a given scene that are completely undamaged. To maintain efficiency, the first step is to identify candidate buildings for damage assessment, using two rules. The first rule involves analyzing the angles, θ, between the surface normals of building points and a horizontal zenith vector (*i.e.*, corresponding to a normal vector of a completely vertical surface). Low values of θ (*i.e.*, corresponding to highly inclined surfaces) have been shown to be indicative of building damage, in both flat and inclined roofs (Aixia *et al.*, 2016). An angle threshold, T_θ, is used to find potential damage points. If a building has a ratio of points that are below T_θ to total points greater than a threshold, T_d, then that building is marked as a candidate for damage assessment. Figure 5.6 shows point clouds for an undamaged building and a damaged building colored by θ. The damaged building had several values in the 40-70^o range, while the undamaged building points were almost entirely 85^o or higher.

The second rule is based on the assumption that damage to buildings often results in debris and portions of the building located at low points around the base of the building. A majority of these points are segmented with the rest of the building points during region growing. Low points are identified as points with a height, h, (taken from the nDSM) below a threshold, T_h. If a building has a ratio of low points to total points above a certain threshold, T_l, then that building is marked as a candidate for damage assessment. Figure 5.7 shows point clouds for an undamaged building and a damaged building, colored by h. Points in the undamaged building range from about 9-13m above ground, whereas points in the damaged building range from 0-2.5m above ground.

Equation 5.8 shows the mathematical notation of the two rules:

$$C(i) = \begin{cases} 1 & if \quad \frac{\#(\theta_i < T_\theta)}{\#\theta_i} < T_d \quad \vee \quad \frac{\#(h_i < T_h)}{\#h_i} < T_l \\ 0 & otherwise \end{cases} \tag{5.8}$$

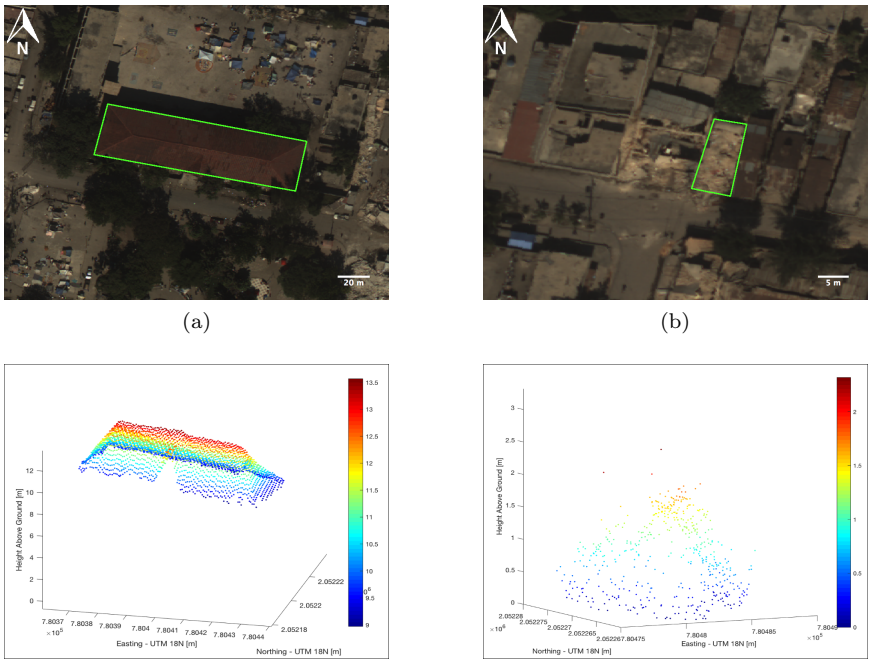
where C(i) indicates if building i is a candidate for damage assessment (a value of 1 indicates a candidate).



(a) (b)

(c) (d)

Figure 5.6: Surface normal angles (θ) of an undamaged and a damaged building in Port-au-Prince: (a) A WASP image of the undamaged building; (b) A WASP image of the damaged building; (c) The undamaged building point cloud, colored by θ; and (d) The damaged building point cloud, colored by θ.



(a) (b)

(c) (d)

Figure 5.7: Heights (h) of an undamaged and a damaged building in Port-au-Prince: (a) A WASP image of the undamaged building; (b) A WASP image of the damaged building; (c) The undamaged building point cloud, colored by h; and (d) The damaged building point cloud, colored by h.

The values of the thresholds will be described in Section 5.5.

Each candidate building undergoes evaluation to classify it as damaged or undamaged. Similar to the candidate identification process, two rules are used for damage classification, a planarity rule and a height rule. An assumption is made that undamaged parts of buildings can be represented as planar segments. The points are segmented into planes using the same region growing algorithm described in Section 5.4.2, but with stricter parameters to ensure planarity ($T_\beta = 4^o$, $T_v = 0.02$). A minimum segment size of 15 points is used to prevent miniature groups of points being counted as true planes. A threshold, T_s, is placed on the ratio of segmented points (those that were grouped into planar segments) to total points (R_s) in a building. If the segmentation ratio is less than T_s, then the building is classified as damaged and removed from the list of candidate buildings. Figure 5.8 shows the segmentation of planes of an undamaged and a damaged building. In Figure 5.8c, almost all of the points are segmented into planes, resulting in a R_s of 0.94 and the building is not classified as damaged. In Figure 5.8d, a building that has been reduced to mostly a pile of debris yields a R_s of 0.58 and the building is correctly classified as damaged.

In some cases, whole roofs or portions of roofs remain intact, but fall to the ground due to damage to the structures that support the roof. The planarity rule is unable to detect these types of damage, so a second rule, based on height, is used. Low points are identified as points with a height, h, below a threshold, T_{h2}. If a building has a ratio of low points to total points above a certain threshold, T_{l2}, then that building is classified as damaged. Otherwise, the building is classified as undamaged.

$$Dam(i) = \begin{cases} 1 & if \quad R_{si} < T_s \ \vee \ \frac{\#(h_i < T_{h2})}{\#h_i} < T_{l2} \\ 0 & otherwise \end{cases} \tag{5.9}$$

where Dam(i) is the classification of building i.

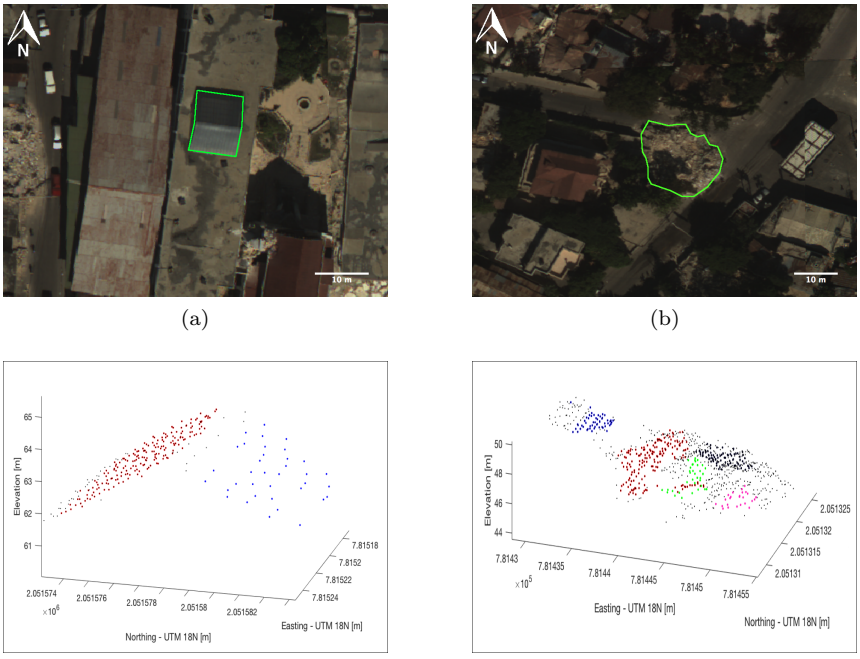


Figure 5.8: Segmentation of planes in undamaged and damaged buildings. Points segmented into planes are colored based on the planar segment, and black points are points that were unsegmented: (a) A WASP image of the undamaged building; (b) A WASP image of the damaged building; (c) An undamaged building is segmented into two planes and very few points are left unsegmented; and (d) A damaged building is segmented into five planes and almost half of the points remain unsegmented.

5.4.4 Experimental Validation

5.4.4.1 Building Segmentation and Damage Assessment Reference Data

In the days and weeks following the Haiti earthquake on January 2, 2010, several organizations contributed to what became an international humanitarian relief effort. A joint collaboration between the UNOSAT, European Commission Joint Research Centre (EC JRC), and the World Bank (WB) generated building damage assessments for most of the affected regions of Haiti. The assessments were made by manually interpreting pre- and post-event airborne (15cm) and satellite (50cm) imagery. The European Macroseismic Scale (EMS-98) (Grünthal, 1998) provides a building damage classification scale from grades I-V, in increasing order of damage, and is frequently used with remotely sensed imagery (Pham *et al.*, 2014). The Haiti damage assessments were classified into four damage categories, roughly corresponding to damage grades I, III-V from the EMS-98: no visible damage, moderately damaged, severely damaged, and destroyed (UNITAR/UNOSAT, 2010). Figure 5.9 shows illustrations from the EMS-98 that represent the damage categories for masonry and concrete buildings used in assessing the Haiti imagery.

A GIS file containing the building damage assessments by the UNOSAT-EC JRC-WB team in point form was obtained and cropped to the seven test sites described in Section 5.3. Building outlines, corresponding to the validation assessment points, were manually traced using 15cm WASP imagery and attributed with the damage grades from the assessment. These building outlines with associated damage grades served as the reference data for validating the building segmentation and damage classifications of the proposed algorithm.

5.4.4.2 Building Segmentation

The workflow proposed in this chapter was applied to the point clouds from the seven test sites from Port-au-Prince and Carrefour. The point clouds were automatically preprocessed and segmented into individual buildings. The 2-D boundaries of the segmented building point clouds were compared with the reference building outlines. Each reference building polygon was labeled a TP if it overlapped with the boundary of a building segmented by the algorithm, or a FN if there was no overlap. Buildings segmented by the algorithm that did not overlap any of the reference polygons were classified as a FP. True negatives were not

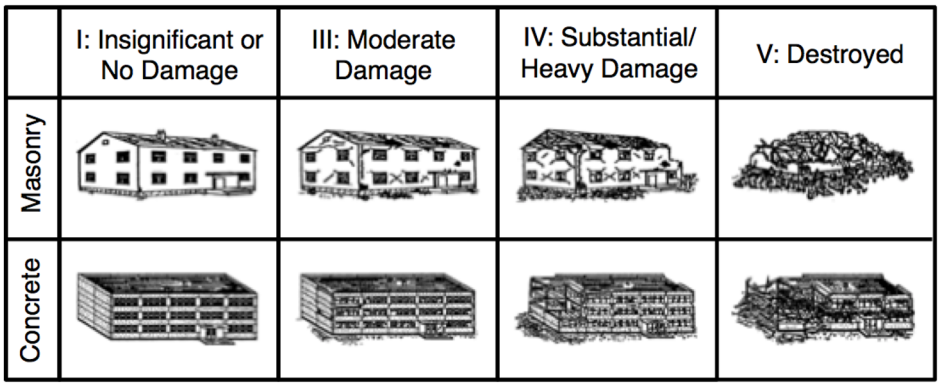


Figure 5.9: Illustrations of the damage grades used for masonry and concrete buildings in the Haiti building damage assessment. Figure adapted from Grünthal (1998).

considered, because all segmented objects are assumed to be buildings. The TP, FP, and FN counts were used to characterize the performance in terms of completeness, correctness, and quality. *Completeness* reflects the percentage of validation buildings that were detected by the algorithm (equation 5.10). *Correctness* reflects the percentage of segmented buildings that were true buildings (equation 5.11). *Quality* is a measure of overall performance that takes into account both the completeness and the correctness of the results (equation 5.12) (Heipke *et al.*, 1997):

$$Completeness = \frac{TP}{TP + FN} \times 100, \tag{5.10}$$

$$Correctness = \frac{TP}{TP + FP} \times 100, \tag{5.11}$$

$$Quality = \frac{TP}{TP + FP + FN} \times 100 \tag{5.12}$$

5.4.4.3 Building Damage Assessment

In order to accurately evaluate the building damage assessment algorithm, it was necessary to de-couple it from the effects of building segmentation. Rather than using the building point clouds from building segmentation, building point clouds were segmented for each building by extracting all of the points from

the preprocessed point cloud that fell within the validation building boundaries. The building damage assessment algorithm was applied to each building, resulting in a classification of damaged or undamaged. If a damaged reference building was correctly classified as damaged by the algorithm, it was labeled a TP. If an undamaged reference building was correctly classified as undamaged by the algorithm, it was labeled a true negative (TN). If a damaged reference building was misclassified as damaged by the algorithm, it was labeled a FN. If an undamaged reference building was misclassified as damaged by the algorithm, it was labeled a FP.

5.4.4.4 Damage Severity Estimation

The automatic building damage assessment methodology presented here is a binary classification of undamaged or damaged. During analysis we grouped all buildings that had damage grades of III, IV, or V into a single damage class. The algorithm was designed to be a rapid and fully automatic way of generating preliminary damage maps, which could be used to identify regions of concentrated damage for emergency response decisions. However, it could be useful to refine the damage classification into severity levels. In this section, we present a method for taking the buildings classified as damaged during damage assessment, and assigning them to a grade of III (moderately damaged), IV (heavily damaged), or V (destroyed) as suggested in the EMS-98 (Grünthal, 1998).

The 811 damaged buildings in the reference set, along with their reference damage severity grades, were used for this experiment. Five features were calculated for each building: minimum height above ground ($minH$), mean distance to neighbors (μ_k), plane segment ratio (PSR), low points ratio (LPR), and the standard deviation of normal vector angles (σ_a). The minimum height above ground, $minH$, is simply the lowest point in the nDSM of the building. The mean distance to neighbors, μ_k, is calculated exactly as it was during preprocessing, but for the individual building rather than the entire scene. The plane segment ratio, PSR, is the same ratio that was calculated during building damage assessment: it is the ratio of points including during plane growing to the total number of points in the building. The low points ratio, LPR, is the same ratio that was calculated during building damage assessment: it is the ratio of points below the

height threshold T_h to the total number of points in the building. The standard deviation of normal vector angles, σ_a, is the standard deviation of the angles between the z-components of the normal vectors and a horizontal zenith vector, for every point in the building. These five features were calculated for each of the 811 buildings, and then they were normalized by the max value for all buildings, so that each feature was weighted between 0 and 1. The normalization was performed to prevent features from carrying more weight based on their scale. These five features were essentially the same ones used for building damage assessment, and have been observed to be indicators of damaged buildings in our dataset.

A feedforward neural network was used to train and classify the damaged buildings. The network architecture consisted of an input layer with five nodes corresponding to the five features, a single hidden layer with ten nodes, and an output layer with three nodes corresponding to the three possible damage grades (III, IV, and V). The network was trained using 55-15-30 split of the data, in which 55% of the data was used for the training set, 15% was used for the validation set, and 30% was used for testing. The scaled conjugate gradient backpropagation method was used to learn the appropriate weights and biases for the different nodes during training because it achieves fast convergence with low memory requirements (MathWorks, 2017). The cross-entropy cost function was used to calculate the error between the predicted damage labels and the reference damage labels.

5.5 Results and Discussion

5.5.1 Building Segmentation Results

Table 5.1 presents the segmentation results for all seven scenes in terms of completeness, correctness, and quality. A visual representation of the segmentation results for scene one are shown in Figure 5.10. The results are overlaid on a WASP image of the scene in Port-au-Prince. The building outlines drawn in green represent TPs, those in red are FPs, and those in blue are FNs.

The algorithm achieved high accuracy for building segmentation, obtaining completeness and correctness scores of over 96%, and an overall quality of 93.46%. Out of 1,953 validation buildings, 1,890 were correctly

Figure 5.10: Building segmentation results for a scene in Port-au-Prince. True positives are outlined in green, false positives are outlined in red, and false negatives are outlined in blue.

Table 5.1: Building segmentation results for all seven of the scenes.

Site #	**# Validation Buildings**	**TP**	**FP**	**FN**	**Completeness [%]**	**Correctness [%]**	**Quality [%]**
1	479	474	20	5	98.96	95.95	94.99
2	243	238	20	5	97.94	92.25	90.49
3	132	129	1	3	97.73	99.23	96.99
4	232	223	4	9	96.12	98.24	94.49
5	326	303	6	23	92.94	98.06	91.27
6	199	191	4	8	95.98	97.95	94.09
7	342	332	8	10	96.03	96.80	93.08
Total	1953	1890	63	63	**96.77**	**96.77**	**93.75**

segmented with only 63 errors of commission and 63 errors of omission. Although the algorithm showed robust performance during the building segmentation stage, it was still important to investigate the source of the errors of commission and omission. Building segmentation preceded the main objective, damage assessment, and therefore errors during the segmentation stage were propagated throughout the rest of the workflow. The following paragraphs take a closer look at the sources of false positives and negatives during building segmentation.

Overall, obtaining 63 false negatives out of 1,953 buildings was a satisfactory result. Analysis of the false negatives revealed that most of them were undamaged buildings. Although it is better to omit an undamaged building than a damaged building in a disaster response scenario, all errors are important to understand. One of the main causes of buildings being omitted was vegetation covering the roof. Although the lidar penetrated the tree canopies and produced returns below the canopy in some cases, the number of returns on rooftops obscured by vegetation was much lower than those with no obscurations. In cases where the vegetation covered a significant portion of the rooftop, buildings sometimes did not meet the 100 point criteria required for segmentation. An example of a false negative caused by vegetative cover is shown in Figure 5.11. Many of the other false negatives were simply a result of the buildings not having enough points and therefore being rejected. There is a tradeoff between eliminating false positives and introducing false negatives with the region size threshold, and unfortunately, allowing buildings with too little points would introduce too many false positives. The best way to overcome this problem would be to collect lidar data at a higher point density.

In total, 63 errors of commission were observed when evaluating the proposed algorithm. False positives were troublesome when passed into the building damage assessment stage, because the planarity assumptions used for rooftops did not hold for other objects, especially when they were not man-made objects. The largest cause of false positives was vegetation. In some cases, the centers of large, dense canopies were not properly classified during the vegetation removal stage. Multiple iterations of graph cuts with adjusted parameters could potentially be used to improve vegetation removal, but the need for parameter adjustment would result in a loss of automation and therefore was not implemented in our workflow. Relief tents were another

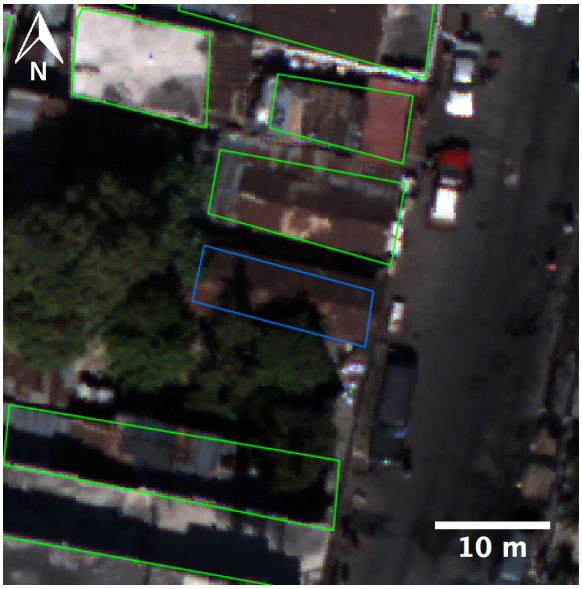


Figure 5.11: An example of an error of omission during building segmentation caused by vegetative cover. The obscuration, coupled with low point density, resulted in the building containing too few returns to be segmented. The omitted building is outlined in blue.

source of errors of commission. The canopy tents were distributed during the emergency response for people displaced by the earthquake. Often times, many of the tents were placed directly next to each other, and as a result, several tents were segmented together as a single building. Figure 5.12 displays an example of multiple canopy tents causing a single false positive. In some cases, groups of two or three tall vehicles in close proximity to each other were mistaken as a building. Despite these sources of errors of commission, a correctness of almost 97% suggests that our algorithm can reliably be used for building segmentation even at relatively low point densities (4.2 pts/m^2 on average).

5.5.2 Building Damage Assessment Results

A total of 1,953 buildings (812 damaged and 1,141 undamaged in the reference data) across seven scenes were classified as damaged or undamaged by the algorithm. The combined confusion matrix for all seven scenes is presented in Table 5.2. An overall accuracy of 78.9% and a Kappa coefficient of 0.57 suggest reasonable

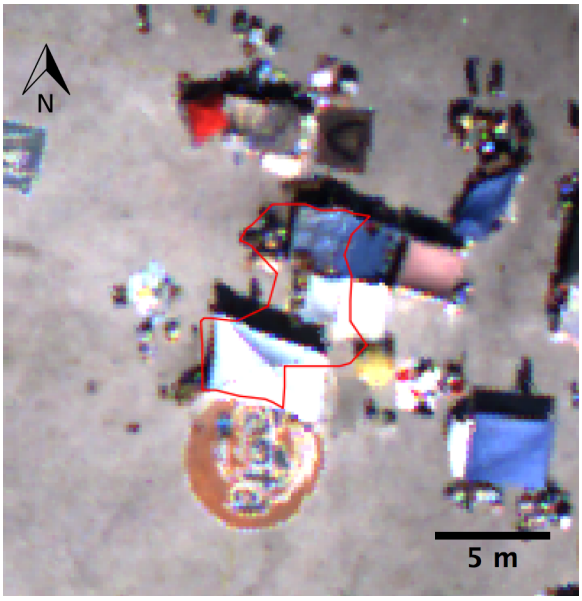


Figure 5.12: Several canopy tents used as makeshift shelters caused a false positive during building segmentation. The false positive is outlined in red.

performance in damage classification, given the wide range of building types and damage patterns in the Haiti scenes. Both errors of commission and omission were caused by several factors, which are discussed in the following paragraphs. A visualization of the results for a scene in Port-au-Prince is shown in Figure 5.13.

In total, 193 errors of omission were observed across the seven scenes. Most of these false negatives were a result of building damage not adequately being represented in the lidar data. For example, in some cases, small portions of a roof would break off and fall to the ground. Although the building damage was

Table 5.2: Confusion matrix for damage classification of all scenes.

		Reference		
	Classification	Undamaged	Damaged	User's Accuracy [%]
Algorithm	Undamaged	922	193	82.69
	Damaged	219	619	73.87
	Producer's Accuracy [%]	80.81	76.23	
	Overall Accuracy [%]	78.90		
	Kappa Coefficient	0.57		

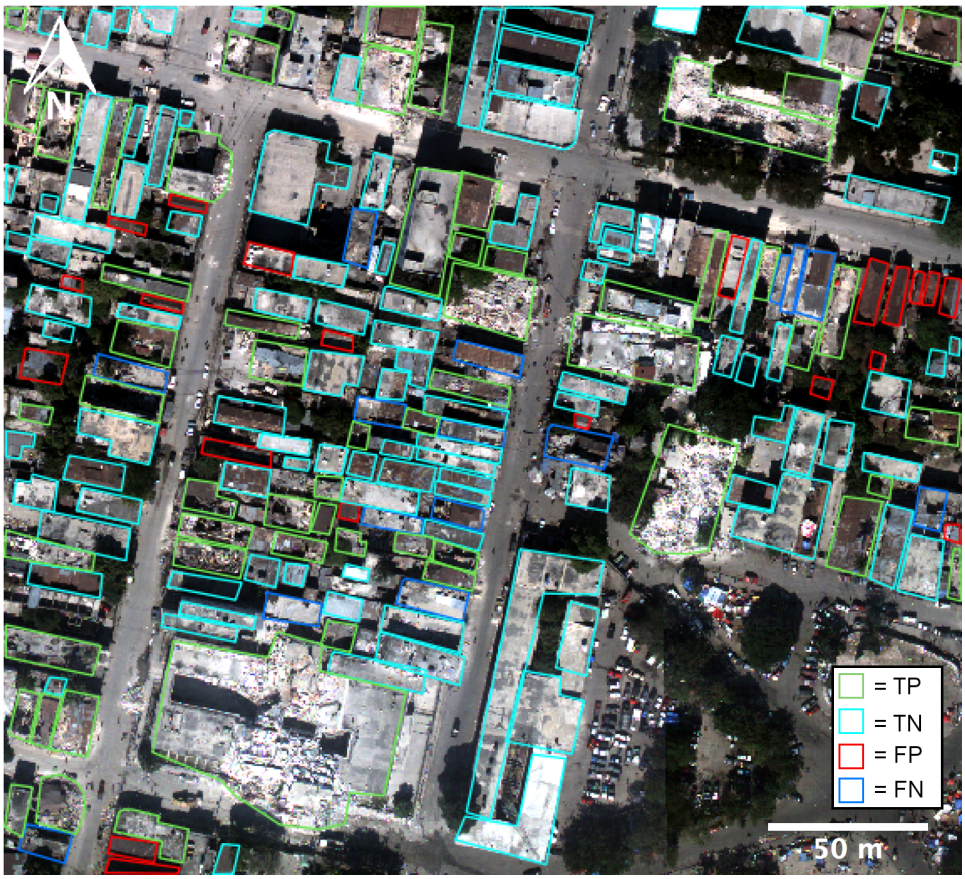
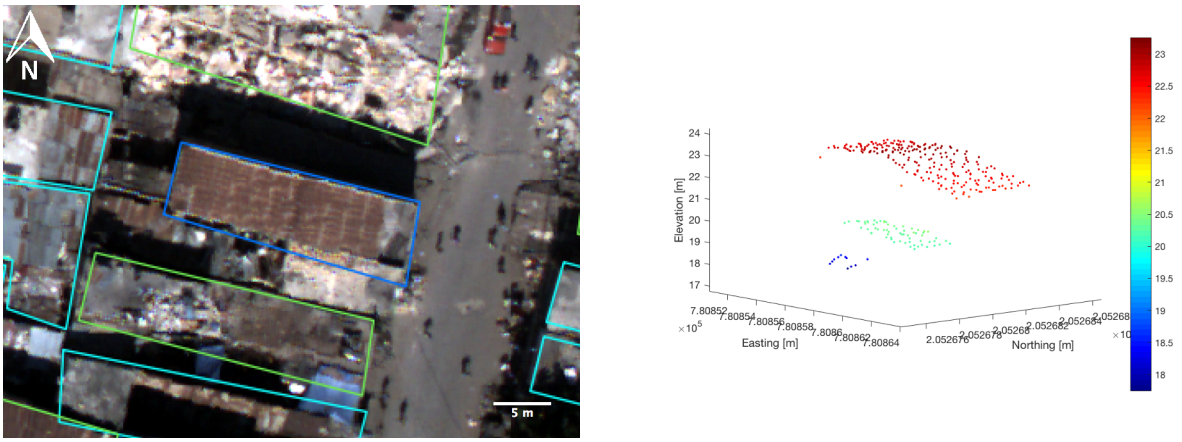


Figure 5.13: Building damage assessment results for a scene in Port-au-Prince. True positives are outlined in green, true negatives are outlined in cyan, false positives are outlined in red, and false negatives are outlined in blue.

visible in imagery, the portions on the ground were low enough to be omitted during preprocessing, and the remaining intact portions of the roof resembled an intact roof. The building damage assessment only considered the intact portion of the roof, and therefore missed the damage. An example of this phenomenon is shown in Figures 5.14a and 5.14b. In some cases buildings that were labeled damaged in the reference set were classified as undamaged by the algorithm, because there was simply no discernible damage in the point cloud or image. Rathje *et al.* (2011) conducted field surveys to assess the accuracy of the the UNOSAT-EC JRC-WB team and found that the assessments were around 77% accurate. The authors even suggested that buildings in grade III should be grouped as undamaged, because of how difficult they are to identify from satellite and airborne imagery, but we chose not to do so because many of the buildings assigned a grade of III were clearly damaged. Despite these issues, 619 out of the 812 damaged buildings in the reference set were correctly classified. At this level of accuracy, the damage maps created by the algorithm would be vital tools for helping direct emergency responders to areas of heavy damage, and dictate regions of interest for detailed damage analysis by ground crews.

Errors of commission were slightly more prevalent than errors of omission, with a total of 219 buildings misclassified as damaged. One of the main causes of false positives was vegetation that was not fully removed during preprocessing. All points that fell within the validation building outlines were used for damage assessment, including vegetation. During the actual workflow of the algorithm, vegetation that remains after preprocessing would typically be removed during the region growing stage of building segmentation due to the merging criteria. As a result of the vegetation remaining in the scene, the buildings were sometimes mistaken as damaged by the planarity rule. Figures 5.15a and 5.15b show an example of a false positive caused by a tree being included with an intact rooftop. Another source of errors of commission were oddly shaped roofs or roofs thats contained many small structures. Due to low point density, these rooftop structures were often undersampled and resembled a damaged rooftop that was not planar. Figure 5.15c shows an example where an undamaged building contained several small, cylindrical structures on the roof. These cylindrical structures were undersampled and were mistaken as damaged by the planarity rule during damage assessment. If the point clouds had been collected at higher point density, these errors would likely occur far



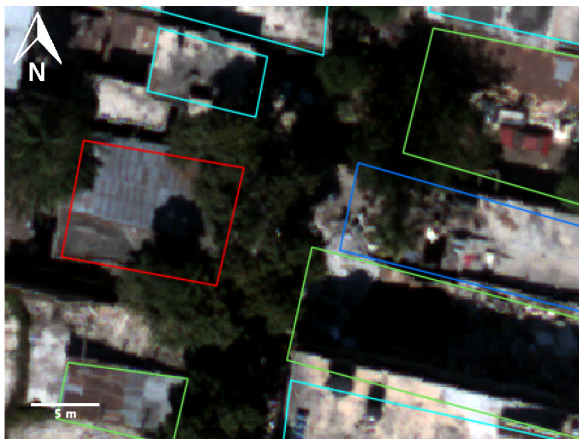
(a) (b)

Figure 5.14: A damaged building is misclassified as undamaged because portions of the rooftop that fell off were too low to be considered, and the remaining roof remains planar: (a) Damage is visible in the image of the omitted building, outlined in blue; and (b) The corresponding point cloud of the building does not adequately represent the extent of the damage.

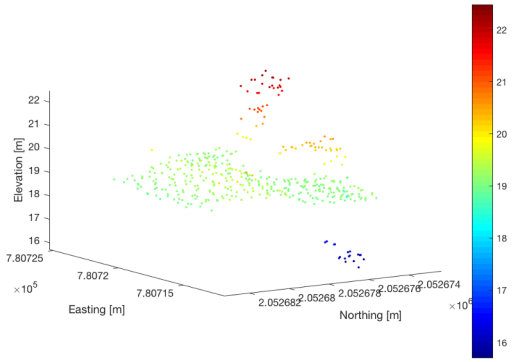
less often. The algorithm still achieved producer's and user's accuracies of 80.8% and 82.7%, respectively, in classifying undamaged buildings, which was reasonable for a preliminary damage map in such a diverse environment.

5.5.3 Damage Severity Estimation Results

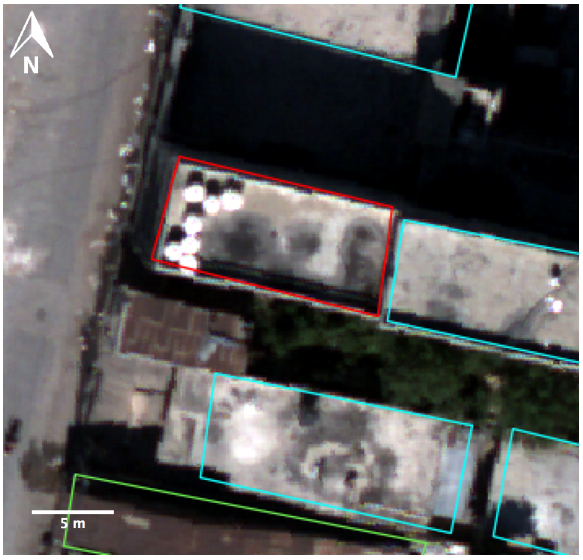
The results of the trained network being applied to the 30% test set, which consisted of 243 buildings, are summarized in Table 5.3. With an overall accuracy of only 66.67% and a Kappa coefficient of 0.39, the results indicate that the network had difficulty discerning between classes. The highest accuracies were obtained for the destroyed class (V) with producer's and user's accuracies of 75.61% and 70.46%, respectively. It was clear that the network tended to overestimate damage, with 12 and 8 class III buildings being mistaken for classes IV, and V, respectively. The network misclassified buildings IV and V almost equally, with 31 class IV buildings misclassified as V, and 30 class V buildings mistaken as class IV. These results are somewhat unsurprising, as class IV and V both represented buildings that had undergone serious damage.



(a)



(b)



(c)

Figure 5.15: Examples of false positives produced during building damage assessment: (a) A WASP image of a building that was misclassified as damaged, outlined in red; (b) The corresponding point cloud of the building in (a). Vegetation is clearly visible in the point cloud; and (c) A building that was misclassified as damaged due to several cylindrical structures on the roof that were undersampled and resembled damage. The building is outlined in red.

Table 5.3: Confusion matrix for damage severity prediction for damaged buildings (classes III, IV, and V).

		Reference			
	Damage Grade	III	IV	V	User's Accuracy [%]
Algorithm	III	4	0	0	100.00
	IV	12	65	30	60.75
	V	8	31	93	70.45
	Producer's Accuracy [%]	16.67	67.71	75.61	
	Overall Accuracy [%]	66.67			
	Kappa Coefficient	0.39			

Another experiment was performed in which buildings in damage classes IV and V were grouped into a single class (referred to as class V for this experiment). Therefore,we were testing to see if the network could distinguish between moderate and serious damage. The same network architecture and parameters were used, except that the output layer contained only two nodes: III and V. The results of this experiment are presented in Table 5.4. The binary classification into a moderately damaged and heavily damaged class performed better than the three-class classification. An overall accuracy of approximately 75% was achieved for the 243 test cases. Moderately damaged buildings were almost twice as likely to be mistaken for severely damaged buildings, than severely damaged buildings being misclassified as moderately damaged buildings. The accuracy of 75% suggests that although the features were able to correctly discern between moderate and severe damage for most buildings, the five features used do not completely capture the damage levels. It is suggested that in a real disaster scenario, at least for the preliminary damage assessment, the automatic method presented in Section 5.4.4.3 is used. However, if a finer classification of damage severity is desired, the neural network method could be applied at the cost of manually creating training data.

Table 5.4: Confusion matrix for damage severity prediction for damaged buildings (class III and classes IV and V grouped into a single V class).

		Reference		
	Damage Grade	III	V	User's Accuracy [%]
Algorithm	III	18	24	42.86
	V	37	164	81.59
	Producer's Accuracy [%]	48.65	87.23	
	Overall Accuracy [%]	74.90		
	Kappa Coefficient	0.22		

5.6 Conclusions

In this chapter we proposed and evaluated an automated technique for assessing building damage via airborne lidar point clouds. Local surface properties of lidar points were used, along with region growing to cluster the points into individual buildings, and then features such as surface normal angle, planarity, and height above ground, were used to make a classification of damaged or undamaged. The building segmentation method was tested on seven point clouds from the 2010 Haiti earthquake, and achieved a detection accuracy of 93.75% for 1,953 validation buildings. The building damage assessment algorithm was tested on the same 1,953 buildings and obtained an overall damage classification accuracy of 78.9% and a Kappa coefficient of 0.57. The main factors that affected the quality of the building damage assessment were vegetation that was not successfully removed during preprocessing, and undersampling of complex rooftops. The results obtained from this research suggest that automated building damage assessment can be used in place of the traditional manual interpretation of imagery with similar levels of accuracy. Automated damage assessment could significantly reduce the time needed to produce damage maps, ultimately leading to faster and more efficient search and rescue missions and prioritization of resources.

Chapter 6

Conclusions

The research in this dissertation has made an effort towards developing automated workflows for analyzing the status of transportation networks and buildings after natural disasters. Specifically, work was completed in context of three objectives:

1. Assess the ability to use airborne lidar point clouds to automatically detect individual piles of debris in the roadway and estimate their volume.

2. Evaluate the ability to use multispectral, aerial imagery and DEM rasters to automatically detect roadway flooding and estimate inundation depth.

3. Assess the ability to use airborne lidar point clouds to automatically detect buildings and classify them as damaged or undamaged.

In accordance with objective 1, a novel roadway debris detection and volume estimation workflow was developed and tested using airborne lidar point clouds from the Haiti earthquake in 2010. The algorithm used local variations of surface normals, along with region growing and filtering techniques to detect and segment debris piles from the point cloud. The piles were then reconstructed as surfaces using alpha shapes, from which their volumes were estimated. The algorithm was able to score completeness and overall qualities

of 97% and 87%, respectively. The volume estimation method was validated through an experiment using debris piles at construction sites and was shown to have an error of approximately 14% or 5.75m^3 at the density of the Haiti point clouds (4.2 pts/m^2) and a minimum error of 5% or 1.5m^3 with high density clouds from the construction sites (18 pts/m^2). The study revealed that false positives from vehicles were limiting factors in algorithm performance. An investigation into using a consumer-grade UAS for debris detection and volume estimation was also completed using a DJI Phantom Pro 3 Quadcopter (see Appendix C). A debris detection accuracy of 90%, and a volume estimation accuracy of 91% were achieved for the debris piles in the construction site.

Towards objective 2, a technique for automatically classifying flood pixels in the roadways and estimating water depth using airborne imagery and DEM rasters was created. A seed flood pixel was automatically detected and used along with target detection algorithms and morphological processing to find the remainder of the flood pixels. Water depth at flood pixels was calculated by subtracting the land elevation of flood pixels from the water surface elevation, obtained at the land-water border. The roadway flooding detection performance was tested using four band airborne imagery from the Mississippi River flooding in January 2016. The proposed method obtained an overall quality of 96.9%, suggesting that it is an effective method for flood detection. The average error in water depth estimation was estimated to be 0.08m. The same experiment was performed on the same imagery, using only the RGB bands, to test the effectiveness of the algorithm in the absence of NIR imagery. The algorithm achieved a quality score of only 69.21% for roadway flood pixel detection, but obtained a completeness of 96.56%. The average error in water depth estimation was almost twice as high at 0.14m. This result demonstrates that the NIR imagery is of great utility for flood detection and depth estimation.

For objective 3, an end-to-end workflow for building damage assessment using airborne lidar point clouds was designed and evaluated using data from the Haiti earthquake in 2010. Buildings were segmented from the point clouds using local surface properties and region growing, and then binary damage classifications were established, based on features such as surface normal angle, planarity, and height above ground. The building segmentation algorithm achieved an accuracy of 93.75% tested on 1,953 buildings from seven Haiti

point clouds. Using the same buildings for evaluation, the building damage assessment algorithm achieved an overall accuracy of 78.9% with a Kappa coefficient of 0.57. Limiting factors in the workflow included vegetation and relief tents being mistaken as buildings, undersampled intact roofs with complex components being mistaken as damaged, and damaged roofs that remained planar being mistaken as undamaged.

The results obtained by the research show great promise for the use of airborne, remotely sensed data in the automatic and rapid generation of information products relating to the status of the transportation network and buildings immediately after a disaster. Such products could dictate important executive decisions in the emergency response, and ultimately save lives.

Appendix A

Generating Debris Shapefiles

In Chapter 3, Objective 1 was addressed through the introduction of an algorithm that characterized roadway debris piles. The algorithm used airborne lidar point clouds and road network shapefiles as inputs, and produced alpha shapes and volume estimates for individual debris piles as outputs. Although not a specific objective of this dissertation, a significant amount of work was devoted to developing a method to automatically generate debris and road shapefiles from the algorithm outputs.

One of the intended uses of the detected debris piles and their volumes was to support a mathematical model that calculated an optimal access restoration plan (ARP). The model took into account network priority metrics (*i.e.*, population and social cost), points of interest (*i.e.*, entry points for emergency vehicles and hospitals), and the estimated time to restore a road (Veras and van Aardt, 2016). Based on these inputs, the model calculated the optimal route between any two points, as well as the order in which each road should be made accessible to minimize impacts to the population.

The restoration times for each road were inferred from detailed information about each road and each debris pile. For each road, the model required the total volume of debris on the road, as well as the passable width (the maximum unobstructed width) that a vehicle could pass through. For debris piles, the model required a high degree of spatial characterization in order to accurately estimate restoration time. To this end, debris piles were sliced in equal-width increments parallel to the centerline of the road. The length and volume of debris in each slice of the pile were used as inputs to the model. Operating at the slice level rather

than the pile level, the model could take into account the ability to clear small portions of debris piles to open a road, rather than having to clear the entire debris pile.

This chapter describes how the road and debris shapefiles were automatically generated from the outputs of the debris characterization algorithm.

A.1 Road Shapefile

For each road, the two edges that were parallel to the direction of travel (and therefore the centerline) were calculated. These edges were needed to determine the passable width of the road, the distances from debris piles to each edge of the road, as well as the slicing coordinates for each debris pile. The points that comprised a road polygon were typically sparse, so a linear interpolation of these points was performed to create a dense sampling of the road boundaries. Next, the road polygon was split into edges by finding points that were located at sharp changes of direction, or corners. For every point, the angle between the vectors formed between its two adjacent neighbors was calculated using the standard angle between two vectors formula:

$$\theta_a = \cos^{-1}\left(\frac{\vec{ab}\cdot\vec{ac}}{|\vec{ab}||\vec{ac}|}\right) \tag{A.1}$$

where $\vec{ab}$ and $\vec{ac}$ are the vectors surrounding point a, and θ_a is the angle between those vectors. Figure A.1 illustrates the angle calculation concept for road boundary points. Critical points that divide two edges of the road were determined by thresholding the angles calculated with equation A.1. The road was segmented into a series of edges by grouping the points that fell between the critical points. The two main road edges were determined using length and other contextual cues and were segmented to be used in calculating important road and debris metrics. For ease of later calculations, the road edges were sampled such that the number of points on each edge was the same and the matching points were approximately across from each other. Figures A.2a and A.2b show a road segmented into all of its edges and only its two main edges, respectively.

Given the two road edges, all of the road and debris metrics could be calculated. For each road, the

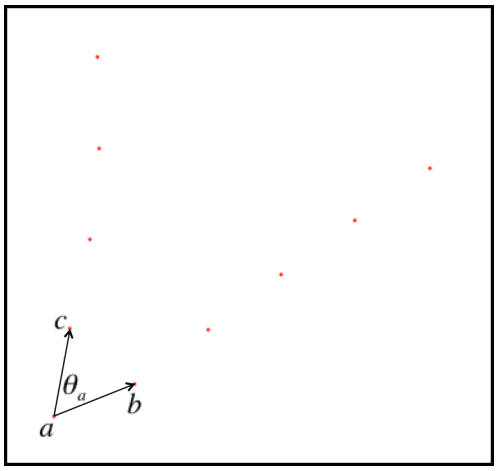
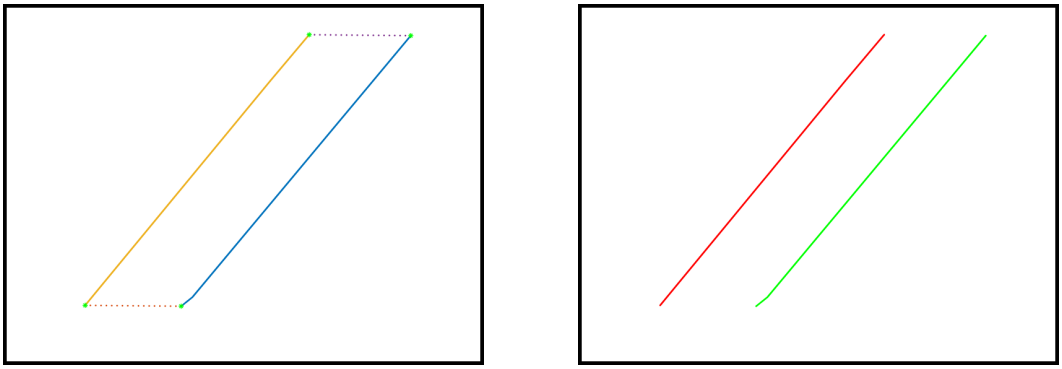


Figure A.1: The angle of each road boundary point is calculated using the vectors between its two neighbors.



(a) (b)

Figure A.2: The two road edges parallel to the direction of travel were extracted: (a) The road boundary was divided into edges using critical points; and (b) The two main edges were segmented for use in deriving debris and road metrics.

mean road width was calculated by simply averaging the distances between each matching point on either side of the road. The passable width was the limiting width a car could pass through, which was calculated as the shortest distance between all of the debris piles and road edges. Figure A.3 shows a road with two debris piles. The unobstructed widths due to the two debris piles, d_1 and d_2, are marked with dotted lines. The passable width of the road is d_1, the limiting distance a car could pass through. The final road metric, the total volume of debris on a road, was simply the sum of the volumes of all of the individual debris piles on the road. In the output shapefile, each road with debris was created as a polygon with the following attributes: the road ID, mean road width, passable width, and total volume.

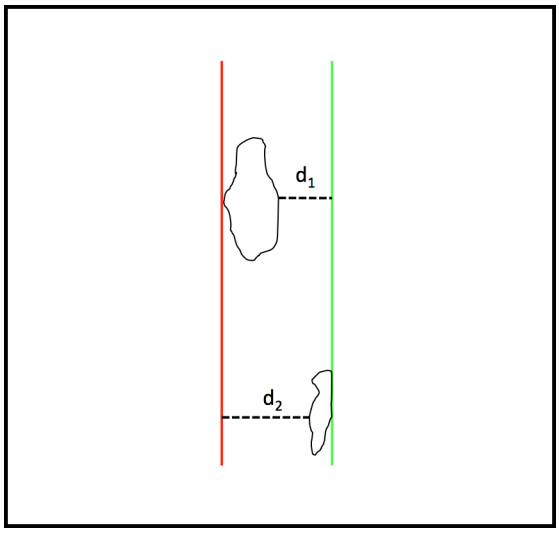


Figure A.3: The passable width of a road is determined by the minimum passable width between the road edge and all of the debris piles. In this illustration, the passable width is equal to d_1.

A.2 Debris Shapefiles

A shapefile containing all of the debris piles as polygons, as well as their attributes, was also created. The distance to each side of the road was calculated as the minimum distance between the debris pile boundaries and each of the road edges. The volume of each debris pile was also listed in the shapefile. In order to

provide a fine level of spatial characterization, the debris pile was sliced parallel to the main two road edges in equal-width increments. Barycentric coordinates were used to densely sample the triangular facets of the alpha shape, providing a much denser, but equal volume, representation of the debris pile. The dense representation was then split by the coordinates of the slices, and then the volume and length of the debris in each slice were calculated and added to the shapefile. The slice metrics were useful in determining if it would be beneficial to clear small portions of the debris in the direction of travel rather than the whole debris pile, to save time. For example, if a debris pile penetrates 2m into the road, it might be sufficient to clear only the 1m farthest from the road edge to allow emergency vehicles to travel on that road. Figure A.4 shows a sliced debris pile and the distances to each road edge.

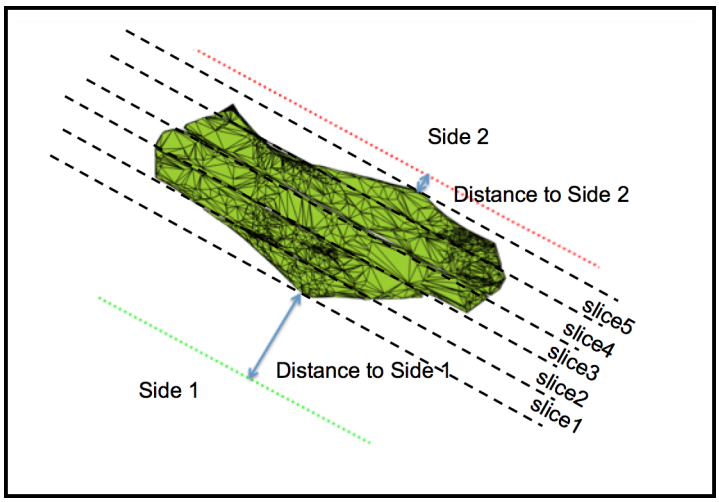


Figure A.4: A debris pile was divided into equal-width slices parallel to the road centerline. Volumes and lengths were calculated for the debris in each slice. The distance to the road sides are shown in this figure as well.

The passable width for each debris pile, which was the maximum of the two distances to the road sides, was also added to the shapefile. If a line connecting corresponding points on either side of the road passed through more than one debris pile, those debris piles were considered as intersecting. The pile ID's and distances between intersecting debris piles were included in the debris shapefile. In the debris shapefile, every debris pile was created as a polygon with the following attributes: the pile ID, the ID of the road the

pile is on, the debris volume, the distance from each side of the road, the volumes and lengths of each slice, and the IDs of and distances to any intersected debris piles.

The road and debris shapefiles were automatically generated and could have been used as information products themselves, or fed into the mathematical model for access restoration. Figure A.5 shows an example road and debris shapefile in an urban area. The shapefiles are colored by road ID, making it easy to see which debris piles belong to each road. The attributes of each road, or debris pile, were stored within the polygons and could be easily retrieved manually or programatically.

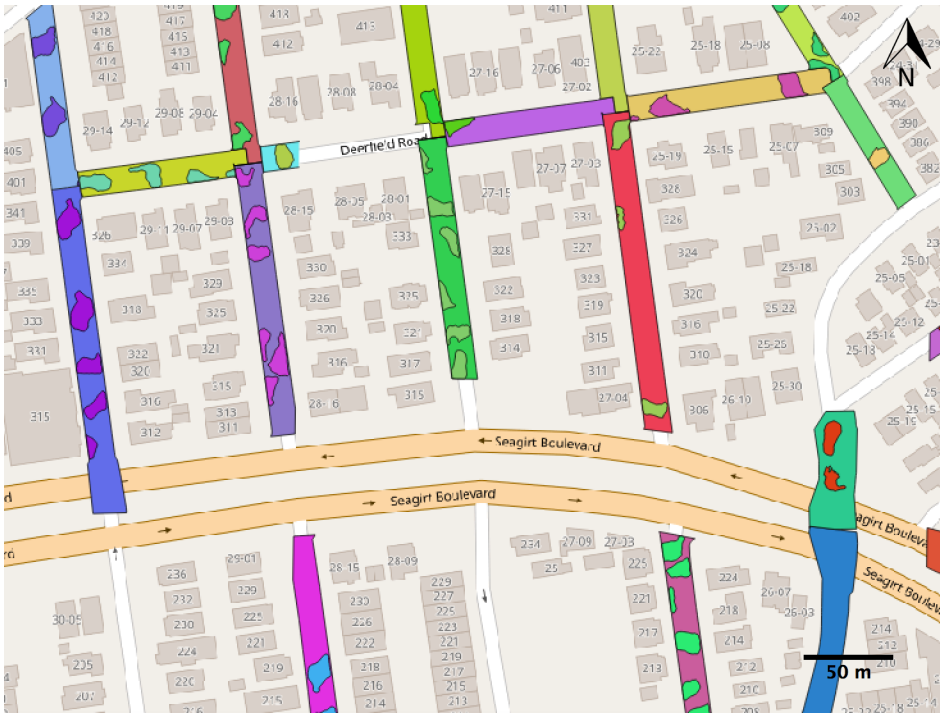


Figure A.5: An example of the road and debris shapefiles produced from the algorithm outputs. Both the road and debris shapefiles are colored by road ID. The shapefiles are overlaid on a street map.

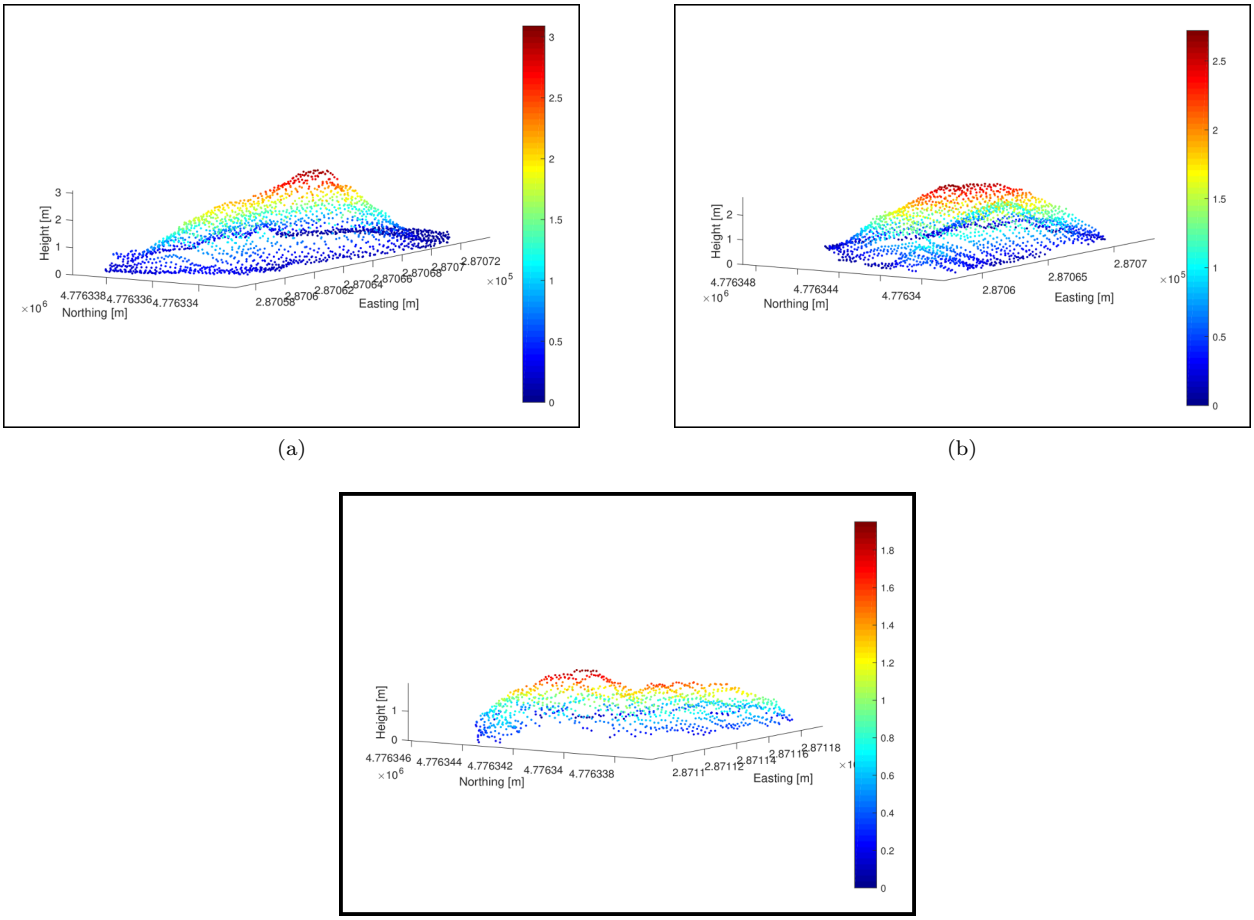
Appendix B

Debris Case Study

Hurricane Sandy was a massive Category 3 storm that struck in October of 2013, causing $75B in damage and over 230 deaths throughout the eight countries in its path (Diakakis *et al.*, 2015). When the storm struck New York City (NYC), USA on October 29, 2013, it generated an estimated 700,000 tons of debris (Gibbs and Holloway, 2013). One region that was hit particularly hard was the Rockaway Peninsula on Long Island. The area is surrounded by water on its north, west, and south borders, leaving it exposed and vulnerable when Hurricane Sandy approached. Thirty-seven blocks suffered severe damage and debris accumulation as a result of the storm. Debris obstructed roadways and forced emergency vehicles to take alternate routes until cleanup crews could remove the debris.

The Rockaway Peninsula was used as a case study to test the debris detection algorithm in an urban environment in the USA, which was one of the intended use cases of the algorithm. No airborne lidar point clouds of adequate point density for our algorithm were collected after the hurricane. However, moderate density (10-12 pts/m^2) point clouds of the area, collected in 2010, were available. The airborne lidar data were collected by Sanborn using a Leica ALS-50 operating at 1,100 m above ground and a pulse rate of 94.7 kHz. Debris piles were artificially placed into the roadways of the lidar data. The set of debris piles that were implanted into the NYC lidar came from the airborne scans of construction sites, as described in Chapter 3. Figure B.1 shows three of the debris pile point clouds that were used for the NYC data.

In order to add debris piles into the roadways of the NYC data, flat areas with no objects were located



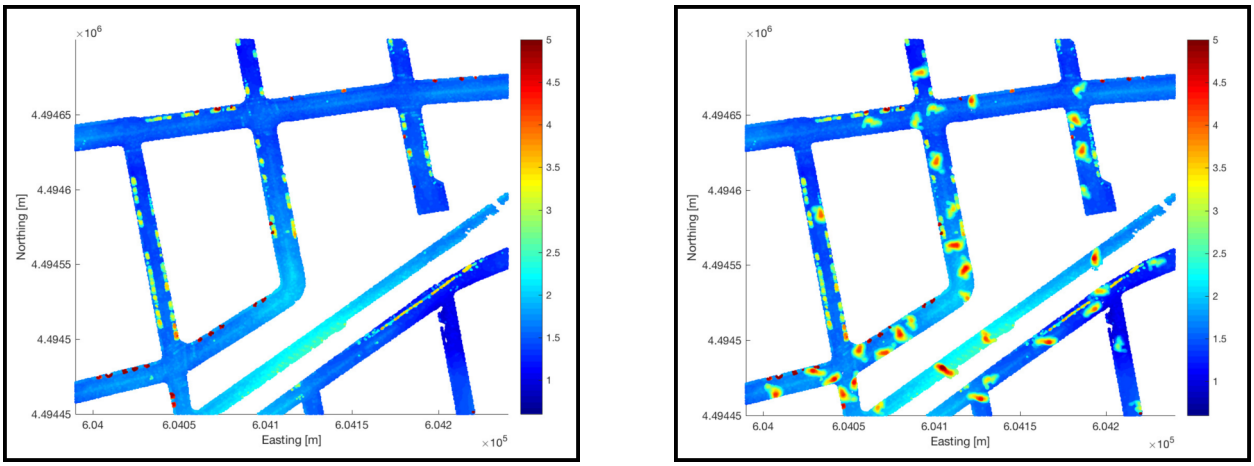
(a)

(b)

(c)

Figure B.1: Three examples of debris piles point clouds that were implanted into the NYC lidar roadways (a-c).

by thresholding the nDSM to only low points. From the candidate low points, a single point was randomly selected. The normal vector of the point was calculated using its 15 nearest neighbors. A plane was fit to the bottom of the debris pile, using least squares fitting, and the normal vector for that plane was obtained using the plane equation. The rotation matrix needed to transform the debris pile normal vector to the scene normal vector was calculated and applied to all of the debris points, forcing the debris pile to be well aligned with the scene points. The rotation step was necessary because the road is not always horizontal, especially in hilly areas. The coordinates of the debris pile were translated to scene coordinates by subtracting the coordinates of the center point from all of the points, and then adding the coordinates of the scene point. Any points that fell within the boundary of the debris pile that were present in the scene originally were deleted, because they would have interfered with the debris surface. This process was repeated until a user-specified number of debris piles had been added to the scene. Figure B.2 shows a point cloud from NYC before and after debris was added.



(a) (b)

Figure B.2: An airborne lidar point cloud of NYC roads before (a) and after (b) the addition of debris.

To mimic the true debris distribution of Hurricane Sandy as closely as possible, Federal Emergency Management Agency (FEMA) Modeling Task Force (MOTF) building damage determination estimates were used. The FEMA MOTF used visible aerial imagery to perform building damage assessments that classified

each building into a damage level of affected, minor, major, or destroyed. The task force also calculated rough estimates of structural debris by assuming 4.1 tons of debris per 1,000 sq. ft. for buildings in the minor damage class and 6.8 tons of debris per 1,000 sq. ft. for buildings in the major damage class (FEMA, 2013). Airborne lidar tiles that contained any buildings of the minor, major, or destroyed building damage class were augmented with debris. Additionally, any tile that was a neighbor of a tile containing buildings with damage, was also augmented with debris. The rest of the tiles were left unchanged. Figure B.3 shows the lidar tile layout for the Rockaway Peninsula with the building damage points overlaid. The tiles in red are tiles that had debris added to them because they contained at least one minor, major, or destroyed building, or they touched a tile that did.

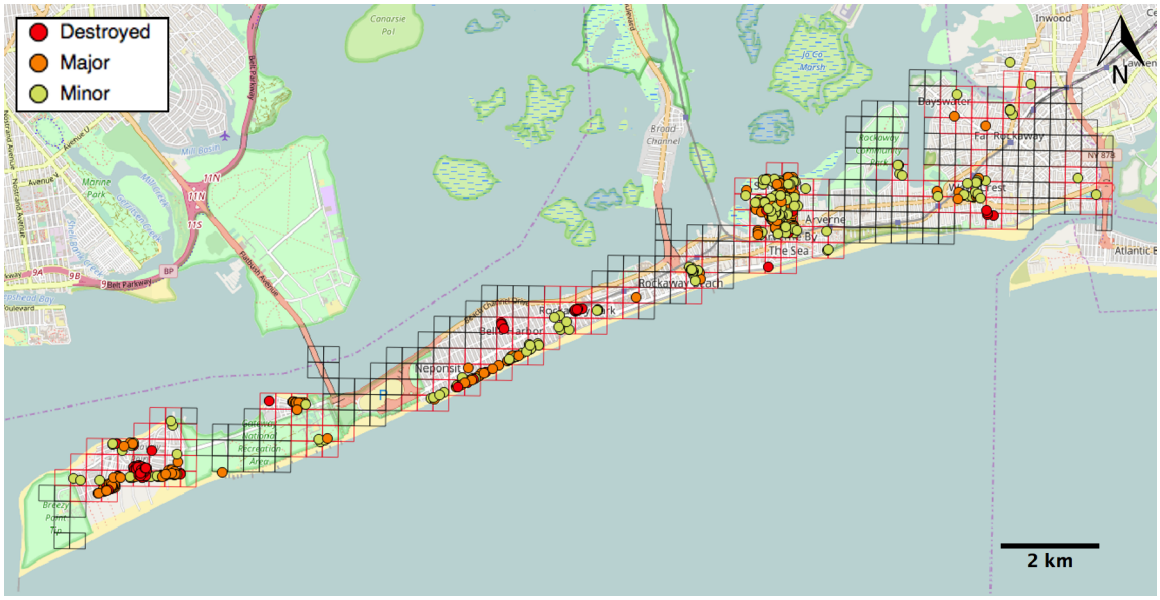


Figure B.3: The NYC lidar tile layout and FEMA MOTF building damage assessments for the Rockaway Peninsula. Tiles outlined with red are tiles that had debris added to them due to the presence of or close proximity to a building with minor or major damage, or a building that was destroyed.

Out of a total of 378 lidar tiles covering the Rockaway Peninsula, 192 had debris added to them. The debris detection and volume estimation algorithm was performed on all of the tiles, and a single road and debris shapefile were automatically produced for the entire scene, using the method described in Appendix

A. Figure B.4 shows the resulting road and debris shapefiles for a portion of the peninsula, colored by road ID. The entire algorithm took approximately two minutes to run per tile, for a total of 12.6 hours for the entire scene. The scene was processed using a personal laptop with a 3GHz Intel Core i7 processor and 8GB of RAM. For a real disaster scenario, the processing could be distributed over many processing cores for much faster processing.

The debris detection results for the case study are presented in Table B.1. Strong detection performance was achieved in the case study, with only 87 of the 1,987 debris piles that were implanted in the scene omitted, resulting in a completeness of over 95%. A total of 32 false positives were detected, leading to a correctness of 98.3% and an overall quality score of 94.1%. Analysis of the results revealed that a majority of the false negatives were caused by debris piles that were smaller than the $20m^3$ volume threshold. After the debris piles were randomly placed in the roads, any points that fell outside of the road network borders were discarded. In some cases the debris piles had significant portions that fell outside of the borders, *i.e.*, that were discarded, despite being above the $20m^3$ threshold prior to being added to the scene. False positives were mainly caused by vehicles parallel-parked on the roadside very close to each other. Despite the false negatives and positives, the case study demonstrated that the debris detection algorithm can automatically generate accurate debris shapefiles for an urban hurricane scenario. The shapefiles can be fed into the mathematical access restoration model to help route emergency vehicles and allocate debris-clearing machinery.

Table B.1: Debris detection results for the Rockaway Peninsula case study.

# Validation Piles	**TP**	**FP**	**FN**	**Completeness [%]**	**Correctness [%]**	**Quality [%]**
1,987	1,900	32	87	95.62	98.34	94.11

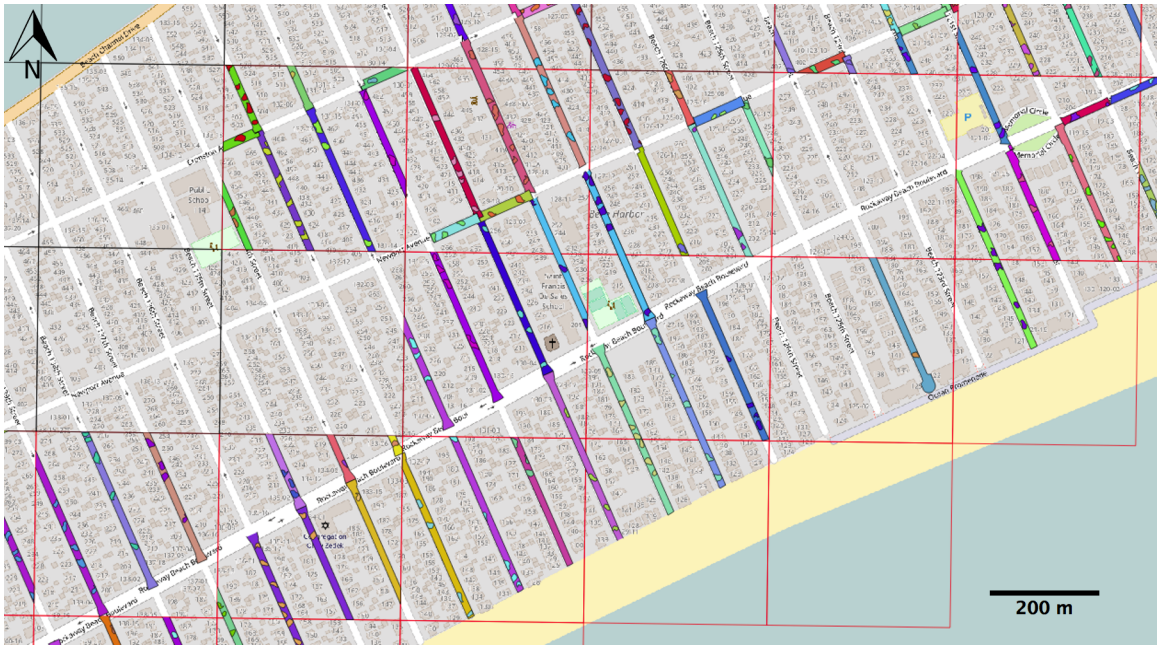


Figure B.4: A portion of the road and debris shapefiles produced for the Rockaway Peninsula case study. The shapefiles are colored by road ID.

Appendix C

Roadway Debris Detection and Volume Estimation with Unmanned Aerial Systems

In recent years, several studies have been published on the use of UASs for collecting post-disaster data (Adams *et al.*, 2012; Yamamoto *et al.*, 2014). Improvements in quality and reductions in cost of these systems have made it feasible to obtain high-resolution imagery of a disaster site with an inexpensive, consumer-grade UAS. Commercial software packages, such as Pix4D and Agisoft PhotoScan, produce point clouds from the imagery using a structure from motion (SFM) processing pipeline (Agisoft, 2016). The UAS point clouds are very dense and contain spectral information, unlike lidar point clouds. Although this dissertation originally sought to assess the ability to use only airborne lidar for roadway debris detection and volume estimation, it was decided during the course of the research to briefly investigate the ability to perform the same task using point clouds collected with an UAS.

A DJI Phantom 3 Professional, consumer-grade UAS was used to capture RGB imagery of several debris piles at a construction site in Rochester, New York. The quadcopter system had a 12.4 megapixel, 1/2.3" CMOS sensor with a 20 mm focal length, operating at approximately 60 m above the ground. The resulting imagery had a GSD of approximately 2.3 cm. A forward and lateral image overlap of 90% was used to ensure several viewpoints of each object in the scene were obtained for subsequent point cloud generation. Figures

C.1a and C.1b display a mosaic of the UAS imagery and an image overlap map of the scene, respectively. The debris piles that were used for debris detection and volume estimation are labeled in Figure C.1a.

Agisoft PhotoScan photogrammetry software was utilized to generate a 3-D point cloud from the UAS imagery. The software used a SFM workflow that involved finding matching points in multiple images, estimating the camera positions, and generating a dense reconstruction of the scene (Agisoft, 2016). PhotoScan produced a high density (82 pts/m^2), colored point cloud of the construction site. The processing time to produce the point cloud from 45 images was approximately two hours using a personal laptop with a 3GHz Intel Core i7 processor and 8GB of RAM. Figure C.2 shows the construction site point cloud generated by PhotoScan.

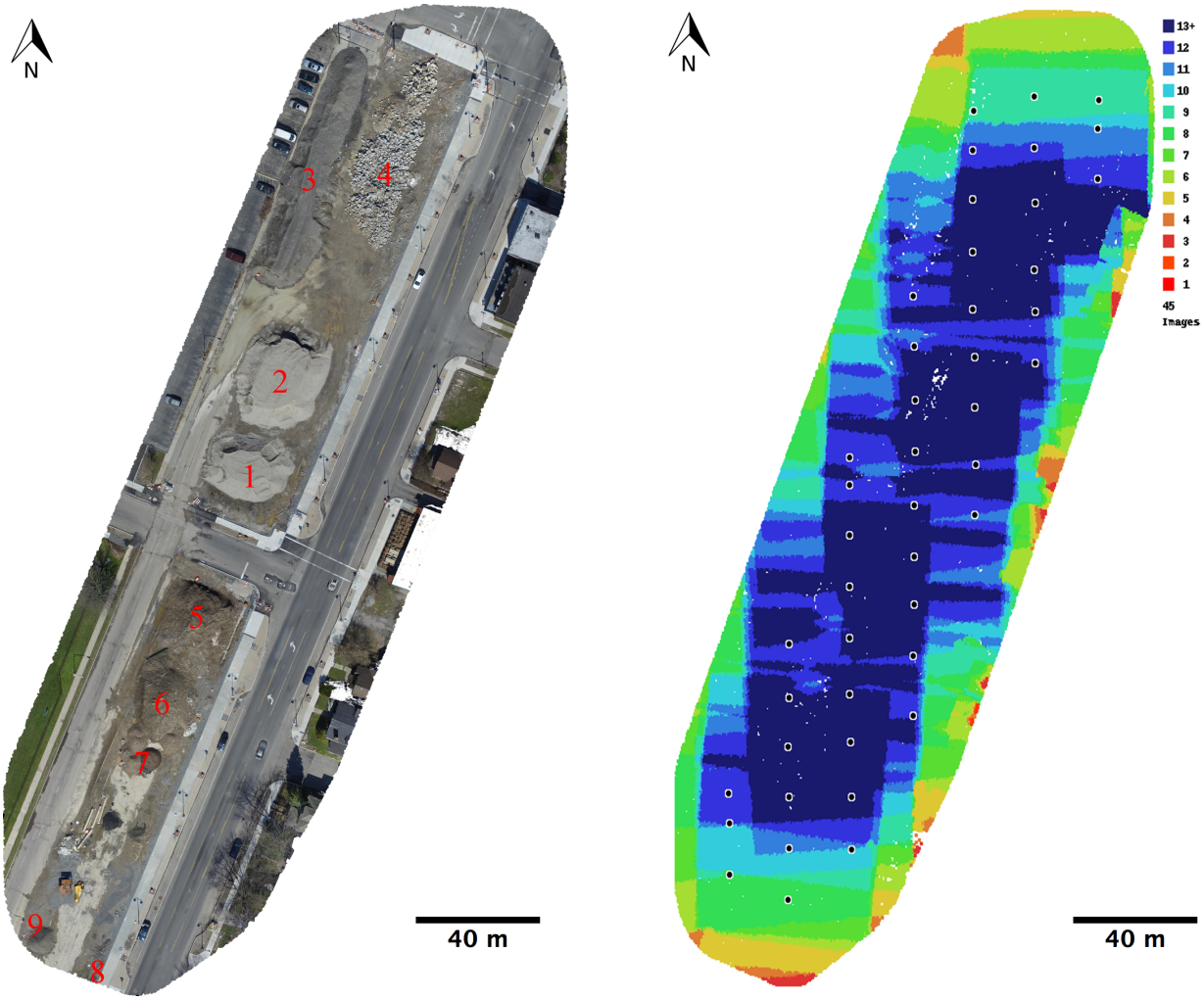
The automatic debris detection and volume estimation algorithm was applied to the construction site point cloud. The construction site was in the median between two roads, so the full point cloud was used as input, instead of the normal input of a point cloud cropped to a road network polygon. The RGB values for the points were ignored for this test.

Table C.1 presents the results of the debris detection. The construction site contained nine debris piles that were over 20 m^3 in volume. The algorithm correctly identified all nine of the debris piles, as well as one false positive caused by construction machinery. As a result, the completeness, correctness, and quality scores for the scene were 100%, 90%, and 90%, respectively. Figure C.3 displays the debris detection results overlaid on a mosaic image of the scene. The boundaries of the debris piles detected by the system that were true positives are shown in blue, and the boundaries of debris piles that were false positives are shown in red.

Table C.1: Debris detection results for the construction site point cloud generated from UAS imagery.

# Validation Piles	**TP**	**FP**	**FN**	**Completeness [%]**	**Correctness [%]**	**Quality [%]**
9	9	1	0	100.00	90.00	90.00

The algorithm automatically estimated the volumes of each detected debris pile using an alpha shape reconstruction. To obtain reference volumes of the debris piles for validation, the volume measurement tool



(a) (b)

Figure C.1: UAS imagery of debris piles at a construction site in Rochester, NY: (a) The RGB imagery and labeled debris piles; and (b) A map of the camera centers and image overlap for the region.

Figure C.2: A colored point point cloud of a construction site containing debris piles. The point cloud was generated using PhotoScan and UAS imagery.

in PhotoScan was used. The tool allowed a user to manually trace the boundaries of an object, and then it estimated the volume using a mesh reconstruction of the region (Agisoft, 2016). Figure C.4 shows the volume estimation approach using the algorithm (extracted point cloud and alpha shape) and PhotoScan (manually tracing the object) for a single debris pile from the construction site.

The volume estimation results obtained for the debris piles in the construction site are presented in Table C.2. For each debris pile, the absolute value of the difference between the measured and calculated debris volumes, in terms of m^3 and %, were computed. The average absolute difference was 52.10 $\pm$ 70.98 m^3, or in terms of percentage, 9.08 $\pm$ 6.79%. The large range in debris pile sizes (1,498.2 m^3 difference between the largest and smallest piles) is the reason that the standard deviation of volume difference is relatively large.

Overall, the debris detection and volume estimation algorithm achieved reasonably accurate results with the UAS-derived point cloud. No reference debris piles were omitted, and only one false positive was detected in the construction scene. The false positive was caused by a folded boom lift with some materials stacked on it, causing it to mimic debris. The colored point cloud of the lift is shown in Figure C.5. Despite the false positive, a quality of 90% and completeness of 100% suggests that point clouds, derived from consumer-grade UAS platforms, can be used for accurate debris detection.

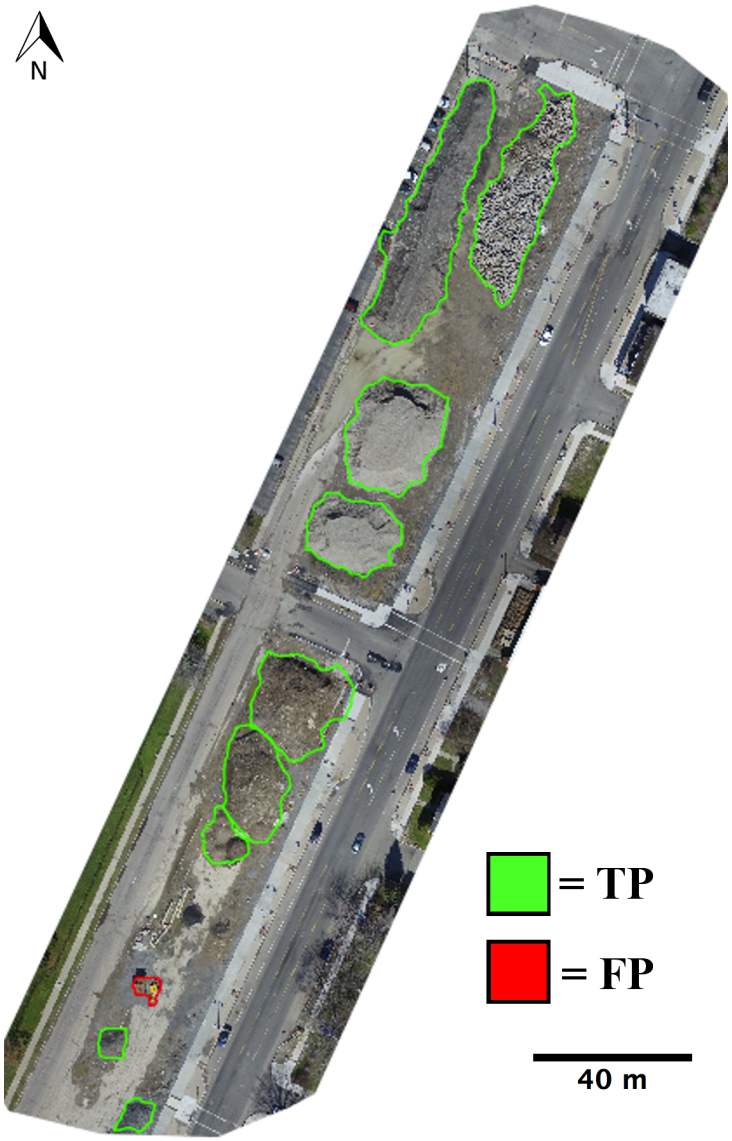
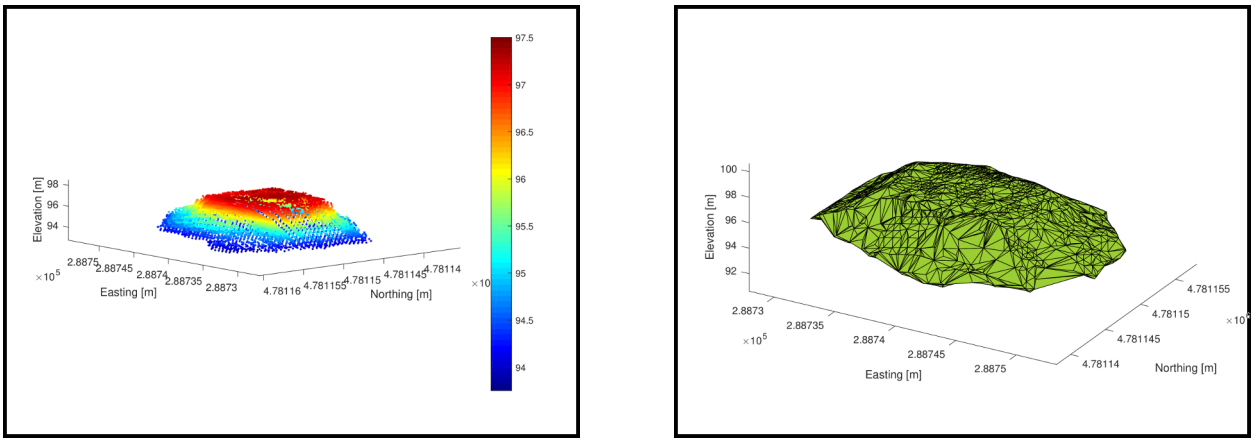
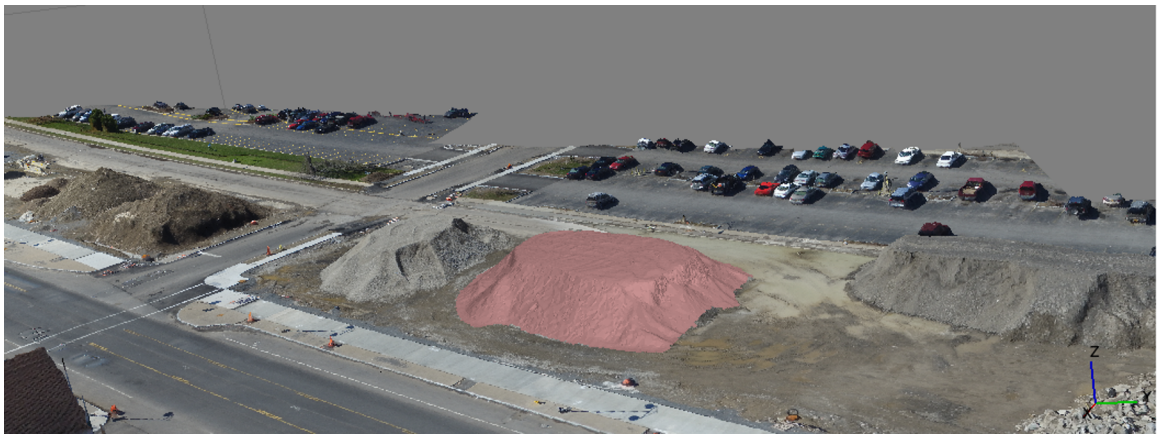


Figure C.3: Debris detection results for the construction site point cloud generated from UAS imagery. Nine out of ten debris piles were correctly detected and one false positive was caused by construction machinery.



(a) (b)



(c)

Figure C.4: Volume estimation for a debris pile from the construction site point cloud generated from UAS imagery: (a) The debris pile point cloud extracted by the algorithm; (b) the alpha shape reconstruction of the debris pile used for volume estimation; and (c) using the volume measurement tool in PhotoScan to calculate the volume of the same debris pile that is shown in (a) and (b).

Table C.2: Volume estimation results for the construction site point cloud generated from UAS imagery.

Pile #	**Measured Volume [m^3]**	**Calculated Volume [m^3]**	**Abs. Difference [m^3]**	**Abs. Difference [%]**
1	572.93	484.66	88.28	15.41
2	1,174.96	1,078.10	96.86	8.24
3	1,530.65	1,741.22	210.57	13.76
4	481.82	480.15	1.67	0.35
5	825.01	768.94	56.06	6.80
6	513.58	511.77	1.81	0.35
7	49.51	45.44	4.07	8.22
8	36.19	38.89	2.69	7.44
9	32.45	25.60	6.85	21.12
		Mean	52.10	9.08
		St. Dev.	70.98	6.79

The error in volume estimation was higher than expected for such a dense point cloud at 9.08 $\pm$ 6.79%. The ability to accurately reconstruct the debris pile surface depends on the level of detail that is captured in the UAS imagery, and therefore is affected by the collection altitude. An altitude of 60 m was used in this experiment to allow the system to safely clear streetlights. Additionally, collecting at a lower altitude requires more images to cover the same area, ultimately increasing the time required for the SFM pipeline. Although an average error of about 9% is slightly high, the results demonstrate that a low-cost ($999), consumer-grade UAS can successfully be used to estimate debris pile volumes.

In a large disaster scenario, debris mapping with a lidar system in an airplane allows for much faster coverage and does not require the computationally expensive process of generating point clouds from images. However, if the region that needs to be mapped is small, or if a lidar system is not available, even a consumer-grade UAS can be used to obtain accurate debris maps and reasonably accurate volume estimates. Although color was not used in this experiment, it could potentially be used during the region growing process or to eliminate false positives.

It should be noted that this experiment was designed to explore the possibility of using a low-cost UAS to map debris. Only nine piles in a single scene were used. A better experiment would involve capturing

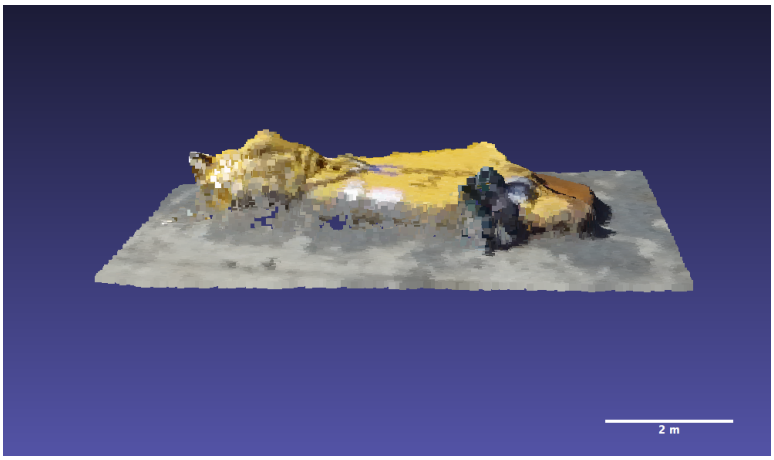


Figure C.5: The folded boom lift that caused a false positive during debris detection at the construction site.

UAS imagery immediately after a natural disaster, ideally with many more than nine debris piles. However, no opportunity to collect such data was available during the course of this research.

Bibliography

Abraham, L. and Sasikumar, M. (2013). "Structural Feature Extraction from Satellite Images". *International Journal of Scientific and Engineering Research*. 4 (12).

Adams, S., Levitan, M., and Friedland, C. (2012). "High resolution imagery collection utilizing unmanned aerial vehicles (UAVs) for post-disaster studies". *Adv. Hurric. Eng.*

Agisoft (2016). *Agisoft Photoscan User Manual*. 1.2. Agisoft.

Aixia, D., Zongjin, M., Shusong, H., and Xiaoqing, W. (2016). "Building Damage Extraction from Post-earthquake Airborne LiDAR Data". *Acta Geologica Sinica (English Edition)*. 90 (4), pp. 1481–1489.

Aksu, D. T. and Ozdamar, L. (2014). "A mathematical model for post-disaster road restoration: Enabling accessibility and evacuation". *Transportation Research Part E: Logistics and Transportation Review*. 61, pp. 56–67.

Ali, S., Hassan, A., Martin, T. C., Hassan, Q. K., Owe, M., Brubaker, K., Ritchie, J., Rango, A., *et al.* (2001). "Geo-spatial tools for monitoring floodplain water dynamics". *IAHS PUBLICATION.*, pp. 465–470.

Amini, J. (2010). "A method for generating floodplain maps using IKONOS images and DEMs". *International Journal of Remote Sensing*. 31 (9), pp. 2441–2456.

Axel, C., van Aardt, J. A., Aros-Vera, F., and Holguín-Veras, J. (2016). "Remote sensing-based detection and quantification of roadway debris following natural disasters". In: *SPIE Defense + Security*. International Society for Optics and Photonics, pp. 98320C–98320C.

Baltsavias, E. P. (1999). "A comparison between photogrammetry and laser scanning". *ISPRS Journal of photogrammetry and Remote Sensing*. 54 (2), pp. 83–94.

Brown, C. (2014). "Waste Management Following Earthquake Disaster". *Encyclopedia of Earthquake Engineering.*, pp. 1–16.

Çelik, M., Ergun, Ö., and Keskinocak, P. (2015). "The post-disaster debris clearance problem under incomplete information". *Operations Research*. 63 (1), pp. 65–85.

Dellepiane, S., De Laurentiis, R., and Giordano, F. (2004). "Coastline extraction from SAR images and a method for the evaluation of the coastline precision". *Pattern Recognition Letters*. 25 (13), pp. 1461–1470.

Diakakis, M., Deligiannakis, G., Katsetsiadou, K., and Lekkas, E. (2015). "Hurricane Sandy mortality in the Caribbean and continental North America". *Disaster Prevention and Management*. 24 (1), pp. 132–148.

Dong, L. and Shan, J. (2013). "A comprehensive review of earthquake-induced building damage detection with remote sensing techniques". *ISPRS Journal of Photogrammetry and Remote Sensing*. 84, pp. 85–99.

Edelsbrunner, H. and Mücke, E. P. (1994). "Three-dimensional alpha shapes". *ACM Transactions on Graphics (TOG)*. 13 (1), pp. 43–72.

Eguchi, R. T., Huyck, C. K., Ghosh, S., and Adams, B. J. (2008). "The application of remote sensing technologies for disaster management". In: *The 14th World Conference on Earthquake Engineering*.

Eismann, M. T. (2012). "Hyperspectral remote sensing". In: SPIE Bellingham.

Elberink, S. O., Shoko, M., Fathi, S. A., and Rutzinger, M. (2011). "Detection of collapsed buildings by classifying segmented airborne laser scanner data". *Proc. Int. Archives Photogramm., Remote Sens., Spatial Inf. Sci.*, pp. 307–312.

FEMA (2010). *Debris Estimating Field Guide*. Tech. rep. 329. Federal Emergency Management Agency.

FEMA (2012). *Preliminary Damage Assessment for Individual Assistance Operations Manual*. Tech. rep. 9327.2-PR. Federal Emergency Management Agency.

FEMA (2013). *Hurricane Sandy Impact Analysis*. Tech. rep. FEMA MOTF.

FEMA (2016). *Damage Assessment Operations Manual*. Tech. rep. Federal Emergency Management Agency.

Feng, C.-M. and Wang, T.-C. (2003). "Highway emergency rehabilitation scheduling in post-earthquake 72 hours". *Journal of the 5th Eastern Asia Society for Transportation Studies*. 5, pp. 3276–3285.

Feng, Q., Liu, J., and Gong, J. (2015). "Urban flood mapping based on unmanned aerial vehicle remote sensing and random forest classifier—A case of Yuyao, China". *Water*. 7 (4), pp. 1437–1455.

Gibbs, L. I. and Holloway, C. F. (2013). *Hurricane Sandy After Action: Report and Recommendations to Maymor Michael R. Bloomberg*. Tech. rep. NYC.

Grünthal, G. (1998). "European Macroseismic Scale 1998 (EMS-98). European Seismological Commission, Subcommission on Engineering Seismology, Working Group Macroseismic Scales. Conseil de l'Europe". *Cahiers du Centre Européen de Géodynamique et de Séismologie*. 15, pp. 15–16.

Guha-Sapir, D., Hoyois, P., and Below, R. (2015). *Annual disaster statistical review 2014: the numbers and trends*. Tech. rep. Brussels: CRED.

Haghighattalab, A., Mohammadzadeh, A., Zoej, M. V., and Taleai, M. (2010). "Post-earthquake road damage assessment using region-based algorithms from high-resolution satellite images". In: *Remote Sensing*. International Society for Optics and Photonics, 78301E–78301E.

Haiyang, Y., Gang, C., and Xiaosan, G. (2010). "Earthquake-collapsed building extraction from LiDAR and aerophotograph based on OBIA". In: *The 2nd International Conference on Information Science and Engineering*. IEEE, pp. 2034–2037.

Hansen, M., Howd, P., Sallenger, A., Wright, C. W., and Lillycrop, J. (2007). "Estimation of Post-Katrina Debris Volume: an Example from Coastal Mississippi". In: *Science and the Storms: the USGS Response to the Hurricanes of 2005*. Ed. by G. Farris, G. Smith, M. Crane, C. Demas, L. Robbins, and D. Lavoie. Vol. 1306. USGS Lafayette Publishing Service Center. Chap. 3, pp. 43–48.

Haralick, R. M. (1979). "Statistical and structural approaches to texture". *Proceedings of the IEEE*. 67 (5), pp. 786–804.

He, M., Zhu, Q., Du, Z., Hu, H., Ding, Y., and Chen, M. (2016). "A 3D shape descriptor based on contour clusters for damaged roof detection using airborne LiDAR point clouds". *Remote Sensing*. 8 (3), p. 189.

Heipke, C., Mayer, H., Wiedemann, C., and Jamet, O. (1997). "Evaluation of automatic road extraction". *International Archives of Photogrammetry and Remote Sensing*. 32 (3 SECT 4W2), pp. 151–160.

Jeyaseelan, A. (2003). "Droughts & floods assessment and monitoring using remote sensing and GIS". In: *Satellite remote sensing and GIS applications in agricultural meteorology*. Vol. 291. World Meteorol. Org. Dehra Dun, India. Geneva, Switz.

Jones, J. L., Haluska, T. L., Kresch, D. L., *et al.* (2001). *Updating Flood Maps Efficiently Using Existing Hydraulic Models, Very-high-accuracy Elevation Data, and a Geographic Information System: A Pilot Study on the Nisqually River, Washington*. US Department of the Interior, US Geological Survey.

Karen, E., Kim, C., Sergey, V., and Vincent, G. (2009). "Remote sensing and the disaster management cycle". *Advances in Geoscience and Remote Sensing", Rijeka: InTech. http://www. intechopen. com/books/advances-ingeoscience-and-remote-sensing/remote-sensingand-the-disaster-management-cycle*.

Kelbe, D., van Aardt, J., Romanczyk, P., van Leeuwen, M., and Cawse-Nicholson, K. (2015). "Single-scan stem reconstruction using low-resolution terrestrial laser scanner data". *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. 8 (7), pp. 3414–3427.

Khoshelham, K., Elberink, S. O., and Xu, S. (2013). "Segment-based classification of damaged building roofs in aerial laser scanning data". *IEEE Geoscience and Remote Sensing Letters*. 10 (5), pp. 1258–1262.

Klemas, V. (2014). "Remote Sensing of Floods and Flood-Prone Areas: An Overview". *Journal of Coastal Research*. 31 (4), pp. 1005–1013.

Korhonen, L., Vauhkonen, J., Virolainenand, A., Hovi, A., and Korpela, I. (2013). "Estimation of tree crown volume from airborne lidar data using computational geometry". *International journal of remote sensing*. 34 (20), pp. 7236–7248.

Kousky, C. (2014). "Informing climate adaptation: A review of the economic costs of natural disasters". *Energy Economics*. 46, pp. 576–592.

Kwan, M.-P. and Ransberger, D. M. (2010). "LiDAR assisted emergency response: Detection of transport network obstructions caused by major disasters". *Computers, Environment and Urban Systems*. 34 (3), pp. 179–188.

Labiak, R. (2011). "A method for detection and quantification of building damage using post-disaster LiDAR data". MA thesis. Rochester Institute of Technology.

Labiak, R. C., Van Aardt, J. A., Bespalov, D., Eychner, D., Wirch, E., and Bischof, H.-P. (2011). "Automated method for detection and quantification of building damage and debris using post-disaster LiDAR data". In: *SPIE Defense, Security, and Sensing*. International Society for Optics and Photonics, 80370F–80370F.

Liu, W., Dong, P., Liu, S., and Liu, J. (2014). "A rich Internet application for automated detection of road blockage in post-disaster scenarios". In: *IOP Conference Series: Earth and Environmental Science*. Vol. 18. 1. IOP Publishing, p. 012124.

Long, S., Fatoyinbo, T. E., and Policelli, F. (2014). "Flood extent mapping for Namibia using change detection and thresholding with SAR". *Environmental Research Letters*. 9 (3), p. 035002.

Malinowski, R., Groom, G., Schwanghart, W., and Heckrath, G. (2015). "Detection and Delineation of Localized Flooding from WorldView-2 Multispectral Data". *Remote Sensing*. 7 (11), pp. 14853–14875.

MathWorks (2017). *Train and Apply Multilayer Neural Networks*. URL: `https://www.mathworks.com/help/nnet/ug/train-and-apply-multilayer-neural-networks.html#bss331l-2`.

Mehrotra, S., Butts, C., Kalashnikov, D., Venkatasubramanian, N., Rao, R. R., Chockalingam, G., Eguchi, R., Adams, B., and Huyck, C. (2003). "Project RESCUE: challenges in responding to the unexpected". In: *Electronic Imaging 2004*. International Society for Optics and Photonics, pp. 179–192.

Messinger, D. (2014). *Remote Sensing Systems and Analysis*. Rochester Institute of Technology Center for Imaging Science.

Messinger, D. W., Van Aardt, J., McKeown, D., Casterline, M., Faulring, J., Raqueño, N., Basener, B., and Velez-Reyes, M. (2010). "High-resolution and LIDAR imaging support to the Haiti earthquake relief effort". In: *SPIE Optical Engineering+ Applications*. International Society for Optics and Photonics, pp. 78120L–78120L.

Murli Gopal Ranjitkar, S. U. (2015). "Post-Earthquake Debris Management: Challenges and Opportunities in Nepal". *Rural Infrastructure*. 6 (6), pp. 71–80.

Musa, Z., Popescu, I., and Mynett, A. (2015). "A review of applications of satellite SAR, optical, altimetry and DEM data for surface water modelling, mapping and parameter estimation". *Hydrology and Earth System Sciences*. 19 (9), pp. 3755–3769.

NASA (2016). *Flooding on the Lower Mississippi*. URL: `http://earthobservatory.nasa.gov/NaturalHazards/view.php?id=87305`.

NOAA (2016). *Billion-Dollar Weather and Climate Disasters: Table of Events*. URL: `https://www.ncdc.noaa.gov/billions/events`.

NOS (2016). *Midwest Flooding Damage Assessment Imagery*. URL: `http://oceanservice.noaa.gov/news/jan16/midwest-flood.html`.

Ogashawara, I., Curtarelli, M. P., and Ferreira, C. M. (2013). "The use of optical remote sensing for mapping flooded areas". *International Journal of Engineering Research and Application*. 3 (5), pp. 1–5.

Otsu, N. (1975). "A threshold selection method from gray-level histograms". *Automatica*. 11 (285-296), pp. 23–27.

Ouzounis, G., Soille, P., and Pesaresi, M. (2011). "Rubble detection from VHR aerial imagery data using differential morphological profiles". In: *Proceedings of the 34th International Symposium Remote Sensing of the Environment, Sydney, NSW, Australia*, pp. 10–14.

Özdamar, L., Aksu, D. T., and Ergüneş, B. (2014). "Coordinating debris cleanup operations in post disaster road networks". *Socio-Economic Planning Sciences*. 48 (4), pp. 249–262.

Padwick, C., Deskevich, M., Pacifici, F., and Smallwood, S. (2010). "WorldView-2 pan-sharpening". In: *Proceedings of the ASPRS 2010 Annual Conference, San Diego, CA, USA*. Vol. 2630.

Pham, T.-T.-H., Apparicio, P., Gomez, C., Weber, C., and Mathon, D. (2014). "Towards a rapid automatic detection of building damage using remote sensing for disaster management: The 2010 Haiti earthquake". *Disaster prevention and management*. 23 (1), pp. 53–66.

Popescu, S. C. (2011). "Lidar Remote Sensing". In: *Advances in environmental remote sensing: sensors, algorithms, and applications*. Ed. by Q. Weng. Boca Raton, FL: CRC Press. Chap. 3, pp. 57–80.

Rabbani, T., Van Den Heuvel, F., and Vosselmann, G. (2006). "Segmentation of point clouds using smoothness constraint". *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*. 36 (5), pp. 248–253.

Ramm, F., Topf, J., and Chilton, S. (2011). *OpenStreetMap: using and enhancing the free map of the world*. UIT Cambridge Cambridge.

Rastiveis, H., Eslamizade, F., and Hosseini-Zirdoo, E. (2015a). "Building Damage Assessment after Earthquake Using Post-Event LiDAR Data". *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*. 40 (1), p. 595.

Rastiveis, H., Hosseini-Zirdoo, E., and Eslamizade, F. (2015b). "Automatic Blocked Roads Assessment after Earthquake Using High Resolution Satellite Imagery". *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*. 40 (1), p. 601.

Rathje, E. M., Bachhuber, J., Dulberg, R., Cox, B. R., Kottke, A., Wood, C., Green, R. A., Olson, S., Wells, D., and Rix, G. (2011). "Damage patterns in Port-au-Prince during the 2010 Haiti earthquake". *Earthquake Spectra*. 27 (S1), S117–S136.

RedCross (2012). *Global Impact Report: Fiscal Year 2011*. Tech. rep. American Red Cross.

Rehor, M. (2007). "Classification of building damage based on laser scanning data". *The Photogrammetric Journal of Finland*. 20 (2), pp. 54–63.

Rehor, M., Bähr, H.-P., Tarsha-Kurdi, F., Landes, T., and Grussenmeyer, P. (2008). "Contribution of two plane detection algorithms to recognition of intact and damaged buildings in lidar data". *The Photogrammetric Record*. 23 (124), pp. 441–456.

Rusu, R. B. (2010). "Semantic 3D object maps for everyday manipulation in human living environments". PhD thesis. KI-Künstliche Intelligenz.

Samadzadegan, F. and Zarrinpanjeh, N. (2008). "Earthquake destruction assessment of urban roads network using satellite imagery and fuzzy inference systems". *The international archives of the photogrammetry, remote sensing and spatial information sciences*. 37 (B8), pp. 409–414.

Sanyal, J. and Lu, X. (2004). "Application of remote sensing in flood management with special reference to monsoon Asia: a review". *Natural Hazards*. 33 (2), pp. 283–301.

Schott, J. R. (2007). *Remote sensing: the image chain approach*. Oxford University Press.

Schweier, C., Markus, M., and Steinle, E. (2004). "Simulation of earthquake caused building damages for the development of fast reconnaissance techniques". *Natural Hazards and Earth System Science*. 4 (2), pp. 285–293.

Shaw, G. A. and Burke, H.-H. K. (2003). "Spectral imaging for remote sensing". *Lincoln Laboratory Journal*. 14 (1), pp. 3–28.

Skakun, S. (2012). "A neural network approach to flood mapping using satellite imagery". *Computing and Informatics*. 29 (6), pp. 1013–1024.

Smith, L. C. (1997). "Satellite remote sensing of river inundation area, stage, and discharge: A review". *Hydrological processes*. 11 (10), pp. 1427–1439.

Stow, D. A., Lippitt, C. D., Coulter, L. L., and Davis, B. A. (2015). "Time-Sensitive Remote Sensing Systems for Post-Hazard Damage Assessment". In: *Time-Sensitive Remote Sensing*. Springer, pp. 13–28.

Sun, S. (2013). "Automatic 3D Building Detection and Modeling from Airborne LiDAR Point Clouds". PhD thesis. Rochester Institute of Technology.

Sun, S. and Salvaggio, C. (2013). "Aerial 3D building detection and modeling from airborne LiDAR point clouds". *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of*. 6 (3), pp. 1440–1449.

Sun, W., Mckeown, D. M., and Messinger, D. W. (2012). "An automated approach to flood mapping". In: *SPIE Remote Sensing*. International Society for Optics and Photonics, pp. 853811–853811.

Sural, S., Qian, G., and Pramanik, S. (2002). "Segmentation and histogram generation using the HSV color space for image retrieval". In: *Image Processing. 2002. Proceedings. 2002 International Conference on*. Vol. 2. IEEE, pp. II–589.

UNITAR (2010). *Density of Bridge and Road Obstacles in Port-au-Prince and Carrefour, Haiti*. URL: `http://www.unitar.org/unosat/node/44/1415`.

UNITAR/UNOSAT (2010). *Joint Remote Sensing Damage Assessment Database*. Tech. rep. UNITAR/UNOSAT, World Bank, and EC Joint Research Centre.

Ural, S., Hussain, E., Kim, K., Fu, C.-S., and Shan, J. (2011). "Building extraction and rubble mapping for city port-au-prince post-2010 earthquake with GeoEye-1 imagery and lidar data". *Photogrammetric Engineering & Remote Sensing*. 77 (10), pp. 1011–1023.

Veras, J. H. and van Aardt, J. (2016). "Decision Support System for Optimal Access Restoration in Post Disaster Environments". Presentation.

Vögtle, T. and Steinle, E. (2004). "Detection and recognition of changes in building geometry derived from multitemporal laserscanning data". *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. 35 (B2), pp. 428–433.

Voigt, S., Schneiderhan, T., Twele, A., Gähler, M., Stein, E., and Mehl, H. (2011). "Rapid damage assessment and situation mapping: learning from the 2010 Haiti earthquake". *Photogrammetric Engineering and Remote Sensing*. 77 (9), pp. 923–931.

Vosselman, G., Gorte, B. G., Sithole, G., and Rabbani, T. (2004). "Recognising structure in laser scanner point clouds". *International archives of photogrammetry, remote sensing and spatial information sciences*. 46 (8), pp. 33–38.

Wang, Y., Colby, J., and Mulcahy, K. (2002). "An efficient method for mapping flood extent in a coastal floodplain using Landsat TM and DEM data". *International Journal of Remote Sensing*. 23 (18), pp. 3681–3696.

Wehr, A. and Lohr, U. (1999). "Airborne laser scanning—an introduction and overview". *ISPRS Journal of Photogrammetry and Remote Sensing*. 54 (2), pp. 68–82.

Xie, H., Luo, X., Xu, X., Pan, H., and Tong, X. (2016). "Automated Subpixel Surface Water Mapping from Heterogeneous Urban Environments Using Landsat 8 OLI Imagery". *Remote Sensing*. 8 (7), p. 584.

Yamamoto, T., Kusumoto, H., and Banjo, K. (2014). "Data Collection System for a Rapid Recovery Work: Using a Digital Photogrammetry and a Small UAV (unmanned aerial vehicle)".

Yao, W., Krzystek, P., and Heurich, M. (2012). "Tree species classification and estimation of stem volume and DBH based on single tree extraction by exploiting airborne full-waveform LiDAR data". *Remote Sensing of Environment*. 123, pp. 368–380.

Yonglin, S., Lixin, W., and Zhi, W. (2010). "Identification of inclined buildings from aerial LIDAR data for disaster management". In: *2010 18th International Conference on Geoinformatics*. IEEE, pp. 1–5.

Zhang, K., Chen, S.-C., Whitman, D., Shyu, M.-L., Yan, J., and Zhang, C. (2003). "A progressive morphological filter for removing nonground measurements from airborne LIDAR data". *IEEE Transactions on Geoscience and Remote Sensing*. 41 (4), pp. 872–882.